EXAMINING THE DIFFERENCES IN READING PERFORMANCE BETWEEN STUDENTS WHO WERE RETAINED VERSUS STRUGGLING READERS PROMOTED IN EARLY GRADES

by

Ashley McClung Potter

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF DISABILITY AND PSYCHOEDUCATIONAL STUDIES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

WITH A MAJOR IN SCHOOL PSYCHOLOGY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2019

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Ashley McClung Potter, titled Examining the Differences in Reading performance Between Students Who Were Retained Versus Struggling Readers Promoted in Early Grades and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Michelle Perfect, Ph.D. Date: 12/17/19

Jennifer Kirkpatrick, PhD Date: 12-17-19

Francesca Lopez, Ph.D. Date: 12-17-19

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Michelle Perfect, Ph.D Date: 12/17/19
Dissertation Committee Chair
Department of Disability and Psychoeducational Studies

ACKNOWLEDGEMENTS

Thank to all of those who have supported during this entire process. I would like to thank Dr. Michelle Perfect for continuing to push me to complete this paper. She spent a lot of time discussing and questioning my analysis to ensure that I knew my material inside and out. I had many trainers and mentors that have not only helped me become the practitioner that I am, but have also have changed the way that I look at intervention data and work with teachers. In particular, thank you to Dr. Debbie Doubrava, Celina LaForge, and Dr. Heather Guynes. I would also like to thank all my work colleagues including fellow school psychologists, teachers, paraprofessionals, administrators, support staff, and all others who work tirelessly everyday in the schools to ensure the well being of our students.

I would also like to thank my family for gently encouraging me to not give up and for knowing when I was tired of hearing, "how's your dissertation going?". I am forever grateful to my husband, Colin Potter, for all the things you did to help me get the time to finish from cooking me dinner, to getting me snacks, and for taking caring of our teething son while I worked. Lastly, graduate school helps you build bonds that are forged in some very stressful situations, so I would also like to thank Caitlan Allen and Sarah Flora. You both have helped me complete this with my sanity intact.

TABLE OF CONTENTS

List of Figures

List of Tables

Abstract

Background and Objectives: Research has shown that retention is not an effective form of intervention and can often delay identification of learning disabilities and increase chances for school dropout. Students can also be retained due to high stakes testing results, most states require students to pass the third grade state reading assessment in order to be promoted to fourth grade. Often students later diagnosed with learning disabilities have been retained. Curriculum based measures (CBMs) are used to help identify reading difficulties at earlier ages for the purposes of interventions. This study compared reading growth, second grade reading scores, and third grade state assessment outcomes between retained and promoted students.

Methods: The current study utilized an existing data set from a school district in southern Arizona that uses CBMs to help identify students for interventions and identification through a Response to Intervention (RTI) process. The sample consisted of 176 students who had scores of <40 letter sounds in a minute on a kindergarten reading CBM. The main dependent variables were second grade oral reading fluency (ORF) score, second grade ORF growth, and third grade reading assessment level. Growth was calculated using a slope formula of spring score minus fall score divided by the number of weeks in between. Independent variables included retention status, special education status, sex, free and reduced lunch at the student's school.

Results: Analyses showed that the retained group (Mdn=22) scored significantly lower on letter sound fluency (LSF) than those promoted (*Mdn*=33), *U*=13.968, *p*<.000. In first grade, those who had been retained (*Mdn*=31) performed significantly lower than those who had not been retained (*Mdn*=49) on their Spring word identification fluency (WIF) score and were more likely to be in the frustrational range (<50 words in a minute) than expected by chance on the Spring administration *U*=10.520, *p*<.001. Further analysis showed that those who were retained

(*Mdn*=33) did not score significantly different on the Fall ORF probe than those not retained (*Mdn*=34), *U*=.269, *p*=.604, two-tailed. Based on a linear regression, no significant differences were observed between the groups for second grade Spring ORF, F (3,172) =.671, p = .571, R^2 = .012. Again using a linear regression, no significant contributions to second grade ORF growth were found, $F(3,162) = 1.63$, p = .185), R^2= .029. Significant unique contributions were made by special education status and Spring ORF, $\chi 2(12, N= 92) = 82.020$, Nagelkerke R^2= .302, p = .004 using a multinomial regression model to determine risk factors for students falling into the Minimally Proficient category on the state assessment.

Conclusion: Significant differences were observed between the retained and promoted groups in kindergarten and 1st grade. No observable differences were observed between the groups in second grade. Retention was not a significant contributor to third grade state assessment category; however, Spring ORF score was. At the end of second grade, 23 (13.06%) out of the original 176 continued to be in the frustrational range. Eleven out of these 23 students scored into the Minimally Proficient category on their state standardized assessment and all but one were identified as receiving special education services. Of the 36 students in the Minimally Proficient category on the state test, 18 were not identified as receiving special education services.

Chapter 1: Introduction

This chapter will cover basic research and perspectives on "struggling" readers. Struggling readers can be defined in multiple ways, but for the purposes of this paper, struggling readers are defined as students who are not reading fluently at their grade level. In order to understand the process of screening and identification, the historical perspective of Specific Learning Disabilities (SLD) and impactful changes in the Individuals with Disabilities Education Act (IDEA) will also be discussed. It will describe the assessment process used as part of response to intervention (RTI). RTI is one approach to early prevention of reading difficulties and identification of SLD. This paper will also explain how the RTI process has been accepted as a systematic intervention to take place within the schools. Finally, this chapter will focus on the mixed or contradictory findings surrounding early retention and whether or not it is beneficial to retain students in kindergarten who are having difficulties with the development of their early reading skills.

Why Target Early Reading Skills?

Despite the necessity of reading for success in school, approximately 65% of fourth grade students are reading below proficiency (National Assessment of Educational Progress, 2013). Proficiency and fluency impact the overall reading comprehension of material (Rasinski, 2004). Research has shown that letter-naming knowledge and phonological awareness are strong predictors of successful reading development in children (Lonigan, Burgess, & Anthony, 2000; Wagner et al., 1997). Further, studies have supported that students who did not read proficiently in the fourth grade had struggled with reading in first and second grades (Juel, 1988; Morris, Shaw, & Perney, 1990). According to the Annie E. Casey Foundation (2012), 63% of students who dropped out of high school had not been proficient readers by the end of third grade.

Developmentally, this time point of end of third grade is important as students have been spending the first few years of school learning how to read. At the end of third grade and the beginning of fourth grade, learning shifts from learning how to read to reading in order to learn and gain information (Anne E. Casey Foundation, 2010).

Differences in reading abilities are evident early in students' schooling years. There are many factors that impact a student's early reading skills. Ortiz et al. (2012) examined several components that could impact early reading including ecological, psychological, and cognitive factors among culturally diverse kindergarten students. Ortiz and colleagues built a model to predict first grade reading performance. The strongest significant predictors of first grade reading performance were letter-word reading and morpho-syntactic skills in kindergarten. Thus, the variables that predicted later reading challenges were reading skills that could be targeted as part of early intervention. Research shows that there are numerous possible reasons for reading skills to be underdeveloped. Reading skills could be limited in kindergarten due to a lack of exposure to reading (Manolitsis, Georgiou, & Tziaki, 2013; McLoyd & Purtell, 2008), a possible disability in reading (Eklund, Torppa, & Lyytinen, 2013), low intellectual functioning (Cunningham & Stanovich, 1997), delayed development (Harris, Botting, Myers, & Dodd, 2011) or low motivation (Fredricks, Blumenfeld, & Paris, 2004; Skinner, Kindermann, & Furrer, 2009). Reading skills could also be hindered because the student has a reading disability that has not yet been identified.

Specific Learning Disabilities

Difficulties in reading often prompt educators to make a referral for special education to determine if the child has an SLD. One difficulty for the school systems is deciphering between students who are struggling readers and students with SLD. Approximately 35% of the special

education population is comprised of students identified under the category of SLD (United States Department of Education, 2015a). Students who are identified as SLD in schools earn lower grades, have higher rates of failure, and have a greater chance of being retained; only 68% leave high school with a diploma (Nation Center for Learning Disabilities, 2014). These negative outcomes for students with SLD can occur for a variety of reasons including: a lack of early intervention, falling too far behind in grade level academics, severity of the disability, and/or a lack of supports (Vaughn & Klinger, 2007; Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007). In order to improve outcomes for students with potential reading disabilities, students who are struggling in reading should be identified early and receive reading interventions to improve their basic skills (Vaughn et al., 2007).

Until 2004, the primary method of determining SLD in school systems was discrepancy model, which assisted in the identification of students with SLD; however, critics have discussed that many of these students were identified too late for remediation through special education to be effective (Fletcher, Coutler, Reschly, & Vaughn, 2004; Fletcher, Francis, Rourke, Shaywtz, & Shaywitz, 1992; Siegel, 1992; Stuebing et al., 2002). The 2004 revision of IDEA included the RTI model as another method to identify children with SLD as well as to incorporate early intervention prior to identification into the school models. RTI strives to assist in the process of intervening with students who struggle with reading and to facilitate early detection in SLD through systematic monitoring of reading skills. Beyond monitoring, RTI provides a structure of interventions for struggling students.

RTI has been found to improve the outcomes for children who might otherwise struggle with reading by providing interventions (Fuchs & Fuchs, 2006; Vaughn et al, 2007). Legislation requires that children learn to read at a proficient level, usually measured on a state assessment.

Legislation varies state to state on what measure is used within the school setting. *Move on When Reading* (AZ 15-701) is an Arizona state law that mandates that all children must be reading at grade level by the end of third grade. Students must be retained if they do not show that they are performing at grade level. There are three exemptions to retention if students score at a falls far below on the assessment; English Language Learners, students who are receiving special education services, or students who are in the special education process and/or diagnosed with a reading impairment are exempt from retention. Although the full extent of this law has not been implemented, between research and the recent laws reading, early intervention, and progress monitoring have become essential to assist a student in academic success.

RTI and Curriculum-Based Measures

As previously mentioned, RTI was developed to assist students who are struggling with reading and give them supports; however, it also identifies students with possible learning disabilities. RTI traditionally has three levels of intervention. The primary level of RTI is a school-wide approach utilizing effective classroom instruction and methods of gathering data from all students to mark their progress overtime. The secondary tier uses evidence-based interventions for students who were identified as having difficulties in the primary level. Students who do not respond to interventions are moved to the tertiary level, which consists of either more individualized interventions or referral for a special education evaluation depending on how a school has decided to structure their framework. This process allows for reading interventions prior to special education referral as many students respond to more intensive reading instruction, but may have otherwise been labeled as SLD or have continued struggling without more intensive remediation. Those who do not respond to reading interventions are referred for a more comprehensive evaluation to better understand the nature of their reading

difficulty (Fletcher et al., 2004; Fuchs & Deshler, 2007; Fuchs, Fuchs, & Compton, 2012; Vaughn & Klinger, 2007). In a typical school, 70-80% of students are successful with regular reading instruction while 20-30% of student will need to undergo some type of intervention. Of those, usually only 10% do not respond to the intervention and may require testing for special education services (Vaughn & Klinger, 2007). Thus, the RTI process reduces inappropriate referrals leaving more time available to provide other types of support. The RTI process helps establish a prevention model for the school system, thus eliminating the need to wait for a student to fall too far below their peers prior to providing additional supports (Fletcher et al., 2004).

The primary method for evaluating growth of reading skills using the RTI model in elementary school is curriculum-based measures (CBMs). CBMs help to guide instruction in the classroom and gather data on effective curriculum. CBMs do not measure the skill or objective that is taught in the classroom, but it measures basic reading skills (Deno, 2003). When examining CBM data there are two important factors to consider. One is the level at which a student's skills are initially measured, for example, this could be measured by counting the number of words students initially read per minute. The second is the slope, which is often referred to as growth in research. In progress monitoring, used in the secondary tier, the skill is measured weekly in order to examine growth. The frequency of gathering data at this level contrasts that of screening at the primary tier, during which the skill is typically measured a few times a school year. By using information regarding both the level and the amount of growth a student has achieved, many important instructional decisions can be made with a fair amount of accuracy. For example, if a student starts with a low level but shows a good trajectory of growth, a steep slope, then the intervention that is being used with that child is working. A student who

has a high level, but flat trajectory is probably reading at a high level and does not need intervention, but could probably benefit from enrichment activities. A student with a low reading level who does not show a trajectory of growth, yielding a flat slope, may need additional supports, such as a more intensified intervention. Data reflecting these patterns are important to identify in schools with a RTI system established in early grade levels (Fletcher et al., 2004, Vaughn & Klinger, 2007; Vaughn et al., 2007).

CBMs are a quick, efficient method to ascertain a student's basic skills and monitor those skills over a period of time. CBMs can be used to improve student and teacher performance as well to screen students who may lack basic skills (Vaughn et al., 2007). CBMs were created to systematically monitor students' progress over time to ensure that teachers were using the same forms to measure student performance over the school year instead of using various forms and mediums to measure growth. There were two systems created for teachers to do this as either a mastery monitoring or as a general outcome measurement. Mastery monitoring is when a teacher teaches a specific skill and then tests on that limited content area whereas general outcome measures tests long-term goals (Tindal, 2013). CBMs created a way for teachers to monitor their instruction effectiveness for individual students by examining the expected general outcome of fluency in basic skills (Deno, 1985; Deno, 2003). This further enabled teachers to monitor the progress of students who had goals to meet under their Individualized Education Plans (IEP; Deno, 1985; Deno 2003; Shapiro, 1989). Screeners are often used as objective measures of a student's reading ability as previous research has shown that teachers have low levels of accuracy when they are asked to predict a student's reading ability (Madelaine & Wheldall, 2005; Ortiz et al., 2012). So by utilizing CBMs, the objective data are used to decide on a course of action versus a subjective teacher opinion.

There are many CBMs that encompass reading, writing, and math with different types for different grade levels. The present paper will only focus on three types of reading CBMs: letter-sound fluency (LSF), word identification fluency (WIF), and oral reading fluency (ORF). LSF is typically used in early grades such as kindergarten to measure many letter sounds a student can accurately identify in a minute. WIF is presented as a list of words that students must read quickly and accurately. ORF is used with students once they have mastered letters and sight words as an indicator of reading skills. ORF is presented as a grade level passage that students read for one minute. The number of words read correctly is calculated to give them an overall score. CBMs will be discussed more in depth in the next chapter.

State and School-Level Policies

In 2010, legislation in Arizona made it mandatory for third grade students to demonstrate proficiency on a standardized reading assessment prior to promotion to the next grade level (AZ 15-211A). Approximately 30 states have adopted legislation that focuses on policies that require standardized reading assessments and interventions to remediate reading deficits. Sixteen states have passed laws that require students to pass a test that demonstrates grade-level reading ability in order to continue to fourth grade (Education Commission of the States, 2014). Although such policies reinforce the importance of reading for future school success, the question is whether retention is an effective intervention to increase growth in reading.

Retention policy and rates. Efforts to target "struggling" readers for reading interventions often begin as early as kindergarten to prevent later reading challenges. The earlier that students are identified as having reading struggles and subsequently provided evidence-based reading intervention, the more likely that they will experience positive outcomes (Vaughn et al., 2007). Kindergarten students who have below expected academic skills or social concerns

are at risk of being retained, particularly in the area of reading (Jimerson & Reynolds, 2003; United States Department of Education, 2006). Retention policies vary greatly across school districts and states. However, the outcomes associated with retention are poor. Students who are retained in elementary school show significantly lower academics by the age of 14 compared to peers (McCoy & Reynolds, 1999).

Jimerson, Anderson, and Whipple (2002) noted that policies and laws have regarded grade retention as a viable or mandated intervention for students who are struggling or fallen behind in the classroom. Research has shown that students are most often held back in early grades (Tingle, Schoeneberger, & Algozzine, 2012; Warren & Saliba, 2012). The United States Department of Education (2006) reported that about 5% of kindergarteners repeat kindergarten even though kindergarten itself is not required in all states (Education Commission of the States, 2016). Warren and Saliba (2012) noted that student retention percentages vary significantly from state to state. Accordingly, the researchers examined first grade retention rates and their results showed that some states have retention rates that are nearly at 0%, whereas others are over 6% with the national average being around 3.5%. Arizona's retention rate is just under 2%.

The rationale behind retention is to give students an additional year to benefit from instruction and repeated coverage of concepts that they have not yet mastered (Jimerson et al., 2002). Identified difficulties in reading may increase the likelihood that a student will be retained as reading is considered fundamental to early academic success; however, grade retention without intensifying interventions or creating a plan of action may lead to another year of unaddressed difficulties and delayed service delivery (Abbott et al., 2010; National Association of School Psychologists (NASP), 2007; NASP, 2011). In fact, one third of students with SLD have been retained at least once (National Center for Learning Disabilities, 2014). The

underlying question is whether using retention as an intervention delays the identification of students who later present as SLD.

Research on retention. The few studies that have shown positive effects for retention have studied the student's retained year or the following year (Jimerson, 2001). Research supports that students make progress during the repeated grade, but those gains are lost over time. Generally, most research has shown that retention does not provide any long-term benefits for students, and it may be detrimental to their education (Hong & Raudenbush, 2005; Hong & Yu, 2007; Jimerson, 2001; Jimerson & Ferguson, 2007; Martin, 2010; Silberglitt, Appleton, Burns, & Jimerson, 2006; Silberglitt, Jimerson, Burns, & Appleton, 2006; Stearns, Moller, Blau, & Potochnick, 2007; Tingle et al., 2012; Warren & Saliba, 2012) and psychological well-being including self-concept (Anderson et al., 2002; Goos et al., 2013; NASP, 2007). Further, the academic achievement of children who were retained has been found to be consistently below that of their new classmates and the students who were retained are unlikely to catch up to their peers (Jimerson & Ferguson, 2007). In this regard, Silberglitt et al. (2006b) found that children who are held back in kindergarten through second grade have been shown to have similar trajectories of reading growth compared to children held back in grades third through fifth. These findings do not support the assumption that holding a child back in earlier grades increases success in subsequent grades.

Perception of retention. Despite research, retained students are often held back based on educator judgments of students' academic progress and on future academic predictions. Range, Holt, Pijanowski, and Young (2012) examined teacher and principal perceptions regarding retention. Their study found that both teachers and principals believed that retention should be used if students do not show adequate academic performance. Teachers and principals reported

that they believed that retention was beneficial to student self-concept in primary grades, in particular, kindergarten. Further, teachers were significantly more likely to report that they believe that retention prevents future academic failure and helps motivate students to attend school. Student perception regarding retention is quite different.

A study by Leontopoulou, Jimerson, and Anderson (2011) had students rank a variety of stressful life events in order of most stressful to least stressful and if they had experienced that life event. They showed that academic retention was the most stressful life event identified by students in the United States even though no students in the study had experienced that event. In fact, retention was ranked as being a more stressful life event than losing a parent or going blind.

Disproportionality in retention. There are disproportional rates of students retained based on certain demographics or individual differences (Anderson et al., 2002). Tingle et al. (2012) conducted a large-scale study to examine characteristics of students who were retained in early grades. This study found that boys were more likely to be retained than girls. African American, Hispanic, American Indian, and multiracial children were also more likely to be retained. Tingle and colleagues further showed that English Language Learners (ELL) and children later identified as special education students were also more likely to be retained than their peers. A study conducted by Gonzalez-Betancor and Lopez-Pulg (2016) in Spain, comprised over 28,000 students. This study showed a strong link between a student's birth date and retention. Those born in the last quarter of birth dates for a particular cohort have double the probability of being retained meaning that those youngest in their grades are most likely to be retained. Other factors that increased likelihood of retention were mother's level of education and father's employment status.

Parental factors have also been shown to significantly impact whether or not a student is retained. Children are more likely to be retained if parents have a lower IQ (Jimerson & Kaufman, 2003). Retention rates also increase with a decrease of parental participation in school as studies have shown that the more parents are involved in school and advocate for their child, retention becomes less likely (Jimerson & Kaufman, 2003; McCoy & Reynolds, 1999). Retention risks are also increased with the number of moves that a child has made (McCoy & Reynolds, 1999).

With regard to outcomes, studies have shown that children who were retained showed more aggression at school (Tingle et al., 2012) and were five times more likely to drop out of school (Jimerson, Anderson, & Whipple, 2002). Hattie (2007) conducted meta-analytic research on various school and home intervention strategies that are often used for students' academic success. Hattie found that retention has a negative impact on student achievement in language arts, math, social studies, reading, work-study skills, and grade point average.

Hong and Raudenbush (2005) examined the effects of retention in an attempt to refute claims that students who were retained were those students who were at risk of poor school performance that predisposed to low academic achievement and future dropout. Hong and Raudenbush examined schools with retention policies for kindergarten and those that promoted all kindergarten students. The researchers created a model for students at-risk for being retained by examining factors that led students to be retained. They applied that model to the schools without a retention policy to find students who would have been identified at other schools as at-risk for retention. Learning was assessed using reading and math US Early Childhood Longitudinal Study Kindergarten (ECLS-K) scores as outcome variables. Their study hypothesized that this finding was due to the students being given repetitive information the next

year instead of being exposed to new information or being exposed to existing information using a different method of instruction. They further concluded that by not exposing students who are retained to new information and concepts, those students have a year of stagnate cognitive development. This study noted that children who had traditionally been retained learned less than if they had been promoted to the next grade level. Results were not significant, possibly due to a lack of statistical power, because only 4.6% of kindergarten students were retained in the schools with retention policies. The researchers found that in the retention schools, students began with high scores in reading and math in the pretreatment year than the non-retention schools. Parents were also more involved. The students they identified as at-risk at the schools with policies of not retaining kindergarteners who were promoted scored higher in overall achievement, reading, and math than their counterparts who were retained at the schools with retention policies. This study found evidence that children who were retained would have learned more had they been promoted as they lost almost a half a year's expected academic growth.

Purpose of the Study

The current study examined the relations between the reading data (LSF and ORF) gathered at multiple time points with students who are retained in kindergarten versus students with similar CBM scores who were promoted into the first grade. The study gathered LSF through second grade ORF data to see differences that exist between those who were retained in kindergarten or first grade and lower in reading and those who were lower in reading and not retained. Additionally, the study uses previous research regarding retention and its possible negative consequences as a starting point to examine if those who were retained have different end of the year reading performance as compared to their counterparts who were promoted. Students identified were retained in either kindergarten or first grade. The students who were

retained have two sets of LSF or first grade ORF scores and have completed the ORF second grade CBMs one year later than those students who were in their same kindergarten cohort but not retained. Results from this study will provide data regarding early reading abilities in children and whether retention is a beneficial intervention for students who show lower reading abilities.

Based on previous research as well as unstudied areas, this study will answer the following research questions:

Research Question 1. Will students who were retained in either kindergarten or first grade demonstrate commensurate reading growth compared to promoted peers at the end of second grade?

Hypothesis 1. Students who were retained in kindergarten or first grade will show commensurate growth on CBMs by second grade compared to students who had similar low reading scores by who were promoted in kindergarten and first grade.

Research Question 2. What are the differences in reading performance between students who were retained in either kindergarten or first grade and their promoted peers at the end of second grade?

Hypothesis 1. Students who were retained in kindergarten or first grade will show commensurate performance on second grade ORF compared to students who had similar low reading scores, but promoted in kindergarten and first grade.

Exploratory Research Question 3. Will students who were retained in kindergarten or first grade be less likely to meet state test requirements for promotion in third grade? At risk for retention will be measured by whether or not students fell into the Minimally Proficient (or equivalent) on the state assessment.

Definition of Terms:

The following terms are used throughout this research study, and based on the literature, can be defined as:

Curriculum Based Measures (CBMs): repeated measures to assess basic academic fluency skills.

Oral Reading Fluency (ORF): a reading CBM that assesses words read correctly per minute on a passage.

Word Identification Fluency (WIF): a reading CBM that assesses how many words a student reads correctly from a list.

Letter Sound Fluency (LSF): a reading CBM that assesses the number of letter sounds read correctly in a minute.

Progress monitoring: process of giving repeated CBM measures over short periods of time in order to monitor growth and make changes in interventions.

Growth: the change that occurs in CBM measures over the course of the school year.

Frustrational range: when the CBM score is below the instructional range threshold for the student's grade level.

Response to Intervention (RTI): is a multi-tiered system of assessments and interventions aimed at specific student needs.

Specific Learning Disability (SLD): is an unexpected underperformance in an academic area. This study is focused on the underperformance in any of the areas of reading.

Special Education: this refers to specialized instruction that is given to students with disabilities who require specialized instruction beyond that provided in a general education classroom.

Minimally Proficient: refers to the lowest score that can be obtained on the Arizona state assessment. Students who score at this level are likely to need additional support to be prepared for the next grade.

Chapter 2: Literature Review

This chapter will cover the history and general framework of RTI. This section will also discuss reading CBMs in more depth and include literature regarding their validity for measuring reading skills. Methods for intervention in relation to RTI and CBMs will be touched upon. A review on the literature regarding retention will be included. This chapter will also discuss other problems that prevent struggling readers from achieving academic success.

Development of Reading

The study of reading development began in the 1930's when researchers first began the study of education (Indrisano & Chall, 1995). Several researchers have developed stages of reading based on readers' achievements and general characteristics (Indrisano & Chall, 1995). Chall developed a model based on reading as a complex task that changes throughout the lifespan development. Chall (1983) described six stages of reading development:

Stage 0: Pre-reading. This stage is also known as pseudo reading occurs between the ages of 6 months to 6 years old. During this stage, the child learns early reading and writing concepts. They pretend to read and they begin to recognize some letters such those in their name. They can also retell stories that have been previously read to them. They learn by being read to and having books provided to them.

Stage 1: Initial reading and decoding. This stage occurs around the ages of 6-7 years old, in grades 1-2. Stage 1 and 2 involve learning how to read. At the beginning of the stage, it is more of a memory game as the child begins to learn about letters and sounds and how to put them together. They can read simple high frequency words and start to use phonetic skills to sound out unfamiliar words. They learn through direct instruction in phonics and repeated practice. They can read simple passages that are at their reading level.

Stage 2: Confirmation and ungluing from print. Children are usually between the ages of 7-8 years old, in grades 2-3. Stage two involves more confirmation of what was learned during stage one rather than learning or gaining new information. The child can read stories that contain familiar sight vocabulary words with increasing speed and accuracy. Children learn through direct instruction in decoding skills and through materials that increase reading fluency. The books read should be familiar to the student in subject matter or structure to help build practice through building fluency. Students should be read to from books above their reading level to further develop vocabulary and language.

Stage 3: Reading for learning the new. This stage has two phases A and B. This stage occurs sometime between the ages of 9-13 years of age, grades 4-8. Phase A is the intermediate stage and phase B is the junior high school phase. This stage begins when children are no longer learning how to read, but reading in order to learn. Texts and passage in this stage and in subsequent stages become more complex in cognitive demands. Reading involves the reading of textbooks and other works that contain new ideas and unfamiliar vocabulary. During phase A, listening comprehension is more important than reading comprehension, but during phase B both are equally as effective for comprehension. Phase A usually has reading materials that introduce new subjects whereas phase B is closer to reading at a basic adult level such as reading newspapers or magazines.

Stage 4: Multiple viewpoints. This stage occurs between the ages of 14-18 years old during high school. Reading involves a wide variety of complex material that have layers of facts, opinions, and concepts. Reading involves a variety of sources and often involves studying words and word parts in order to pull apart meaning. When reading materials are more difficult,

reading comprehension may be better than listening comprehension. Most skills at this stage are developed through reading in a variety of areas usually through formal education.

Stage 5: Construction and reconstruction. This stage involves students 18 years and older in college. Reading is often for personal and professional purposes. Reading can now be used to integrate one's own knowledge with that of others and can assist in creating new knowledge. Reading is done in detail and for a specific purpose. Reading involves difficult materials and assists in writing papers and essays.

Etiology of Struggling Readers

Students may struggle with basic reading skills; however, there are a multitude of underlying factors that could impact the attainment of those skills. A longitudinal study by Cunningham and Stanovich (1997) found that 1st grade reading abilities were a strong predictor of 11th grade outcomes after controlling for cognitive abilities. As this study indicated, early reading difficulties can contribute to lifetime problems in school and academics. These are factors that must be closely examined prior to labeling any child with SLD. Some of these factors can be altered through intervention. If children enter kindergarten with a disadvantage in reading, the gap tends to stay the same or widen over time (McLoyd & Purtell, 2008). These disadvantages that students enter school with could be due to environmental, biological, or motivational factors. A lack of exposure to reading could be attributed to several factors. It could be due to lack of parental involvement in early reading activities, poor instruction in reading, or some type of other educational disadvantage in reading. Manolitsis et al. (2013) found that parents' teaching of reading skills and home exposure to story books increased children's letter knowledge, phonological awareness, reading fluency, and vocabulary skills. Additionally,

engagement in classroom activities and motivation helps to support student's efforts and attention to become better readers (Fredricks et al., 2004; Skinner et al., 2009).

Harris et al. (2011) studied the relations between speech impairments and early reading development. They found that children with delayed speech development performed similarly to typically developing peers on reading measures and showed positive early reading skills. Children with speech disorders struggled on reading measures in tasks with phonological awareness and did not show emergent early reading skills. This may indicate that students with delayed speech may not struggle with reading, whereas students with speech disorders may have a strong impact on their reading skills over time.

There are other factors that contribute to early reading skills that may not be able to be easily remediated. Children who have compromised early cognitive development are at risk for developing a reading disability; however, if these students had a high involvement in the classroom, there tended to be an absence of a reading disability (Eklund et al., 2013). Geva and Massey-Garrison (2012) studied English Language Learners (ELL) and English as first language students who were poor at decoding, comprehending, and average readers. Normal readers performed better on oral language measures than the other two groups. They further found that regardless of the language status, students who had poor comprehension or decoding skills, had difficulties with aspects of language. Students with learning disabilities also struggle to remediate difficulties in reading without intensive intervention. Further, most children with a reading disability have family history of dyslexia which makes early reading exposure at home more difficult (Eklund et al., 2013).

Background of Response to Intervention

IDEA discusses three models for the identification of SLD in the schools. Those models are the discrepancy model, response to intervention (RTI), and other research based methods (IDEA, 2006). Current practices show that many districts use patterns of strengths and weaknesses for SLD identification as an "other research based method". When the identification of SLDs began in the 1970s, the only method used was the discrepancy model that measures the difference between a student's achievement and aptitude and that difference must be an approximately 22 points in order for that student to qualify. Decades after its inception, it became clear that there were empirical and practical disadvantages to standards of practice for SLD identification within the discrepancy model. Schools were concerned that SLD was being over identified costing schools billions of dollars on special education services (Fuchs & Fuchs, 2006). Schools began using different methods to reduce the number of students being labeled as SLD as well as reduce the number of students referred for the special education process as students were being unnecessarily served in special education. Francis et al. (2005) discussed how cut points can be arbitrary when they are used in isolation, such as with the IQ-achievement discrepancy. The discrepancy is not sufficient to identify a child with SLD. There is not strong empirical evidence guiding the decision as to how much of a discrepancy should exist to constitute a SLD. Research has found that special education had too much emphasis on the discrepancy model instead of focusing on strategies that provide prevention and intervention in the school systems (Francis et al., 2005; President's Commission on Excellence, 2002). Further, preventative methods should focus on checking progress of all students to ensure their academic growth as they progress over the years (President's Commission on Excellence, 2002). RTI was used in schools for many years prior to its inclusion in IDEA, but it was not organized into the

system it is now (Brown-Chidsey & Steege, 2011). In 2004, IDEA added RTI and alternative research based methods to the accepted models for the identification of SLD in the schools.

RTI is often referred to as a multi-level approach involving different tiers of supports starting with the population of the school. Francis et al. (2005) related RTI to the field of medicine that gathers data on blood pressure that "establishes a historical record to base decisions on versus a decision on a one-time measure" (p. 105). RTI focuses on both early intervention and early identification by gathering many data points over time. This system of data can be reviewed as part of the student's educational history to better target interventions to increase student learning performance. Early intervention can reduce the number of students who are diagnosed with SLD. The RTI process prevents students who struggle with reading from inadvertently being diagnosed with SLD when in fact the reasons for their compromised reading skills may be a combination of maturation (Harris et al., 2011), lack of exposure (Manolitsis et al., 2013; McLoyd & Purtell, 2008), a disability in reading (Eklund et al., 2013), low cognitive skills (Cunningham & Stanovich, 1997), or low motivation (Fredricks et al., 2004; Skinner et al., 2009).

If reading problems are identified and targeted early enough, students have the opportunity to show growth in reading to close the gap in reading skills compared to same aged peers (Fletcher et al., 2004). RTI is based on the concept of schools using evidence-based interventions with children, monitoring their progress, and making educational decisions based upon the children's response to the interventions used (Fuchs & Deshler, 2007). Unfortunately, according to Speece (2002), teachers often do not make referrals for children to be evaluated for a possible learning disability until the students have struggled for a prolonged time or evidenced a pattern of reading failure. Research has shown that students who struggle with early reading

skills do better when they receive interventions and supports than those who are only provided regular instruction reading supports (Mathes et al., 2005). Evidenced-based interventions focus on increasing the reading levels of students who may be struggling regardless of whether it leads to future diagnosis of SLD. Vellutino, Scanlon, and Lyon (2000) followed students from kindergarten through fourth grade to examine their reading abilities using a variety of standardized achievement measures. The study identified two types of readers: those who had not had proper exposure and/or instruction in reading and easily could be remediated through intervention and those with reading difficulties that were more challenging to remediate as they most probably had an actual disability in reading. Through the process of RTI, both types of readers can be assisted through the process of evidence based interventions regardless of future identification.

RTI was added to IDEA to not only assist in defining and identifying learning disabilities but to develop and implement prevention models that would reduce the prevalence of learning difficulties particularly in reading (Fuchs & Deshler, 2007). Although RTI has strong potential and good scientific evidence, its utility is not fully being realized. Collecting data on every student allows for systematic tracking of student progress over in a longitudinal fashion to make educational decisions. Some of these educational decisions may include the need to intervene with a different instructional model for a short time, to give the student additional classroom supports, or possibly the need for a referral to special education. Using universal screening tools help to correspond decision making between team members in line with the data (Shapiro et al., 2012). However, data are only useful if it is interpreted appropriately to make sound decisions.

Judge and Bell (2011) conducted a study to examine reading growth within different subgroups of students who were identified as having SLD in elementary school over a six year

period of time. The researchers retrospectively analyzed data and separated students who had been identified as having SLD into three groups that were classified as early-emerging SLD, emerging SLD, and late-emerging SLD. Students were classified into the early emerging SLD group if they were identified in kindergarten or first grade, emerging SLD if they were identified in second or third grade, and the late-emerging group if they were identified in fourth or fifth grade. The reading assessment given to the participants included measures of early literacy and passage comprehension. Early literacy focused on phonemic awareness, print familiarity, single-word decoding, and vocabulary, whereas comprehension focused on initial understanding, developing interpretation, person reflection, and critical stance done using sentences, paragraphs, and stories. Researchers found that students who were identified in elementary school as SLD had significantly lower reading achievement at the entry to kindergarten. Initial reading gaps in kindergarten were associated with SES status. Students who were non-African American, non-Hispanic, non-Asian/Pacific Islander, or were non SLD girls had the highest amount of growth per month. Students identified with early emerging, emerging, or late emerging SLD had a reading growth coefficient of -0.35 compared to Caucasian females who had a coefficient of 1.80. This study found that low SES background, being male, minority status, and having SLD were significant predictors related to student reading growth in elementary school. Further, they found that grade retention was not a statistically significant predictor of reading growth. The researchers found that initial reading scores in kindergarten and rate of reading growth in kindergarten tended to continue a higher level of reading growth than students with lower reading achievement. They did not find significant differences between the three types of SLD groups meaning that when students were identified as having SLD, did not impact the amount of growth they showed in reading performance. This study found that growth is more significant in

the general education population, is more substantial in elementary school, and more reading growth occurs in the Fall although those in special education do not show as significant of an impact with this season effect. This study did not address directly address interventions as the study focused on reading growth. The study further suggested that early diagnosis did not impact prognosis in SLD and suggest that the data from this study help evaluate instructional effects and data decision making when examining reading data. The study concludes that early diagnosis should not be the focus of RTI, but rather, the ability to give students appropriate instruction and intervention earlier.

Framework of RTI

Multi-tiered systems of support. RTI typically involves different levels of interventions; most often schools have a three-tiered approach. The first tier is the general instruction that is taught in the classroom. In the first tier, a universal screener is given to systematically gather data on student are progress in the classroom as well as to evaluate the effectiveness of the classroom instruction (Vaughn & Klinger, 2007). In fact, some professionals refer to RTI as "response to instruction" because the interventions are types of instruction that are matched to the learning needs of the child (Brown-Chidsey & Steege, 2011). The second tier is an individualized or group intervention for those students that were low on the screener. The second tier usually takes several weeks until students have had sufficient exposure to the intervention. This intervention is usually done either one-on-one or in small groups. The final tier involves examining the data gathered from the intervention for possible referral for an evaluation process to determine the need for special education services and/or another more intensive intervention (Vaughn & Klinger, 2007).

Universal supports. Tier one is comprised of several elements for an effective RTI process. The first element that is needed is an effective curriculum that is evidence-based (Fuchs & Deshler, 2007). According to Vaughn et al. (2007), the next important element in the first tier is screening and benchmark testing of students at least three times a year. This step is essential as it gathers data on important basic grade-level concepts. The integrity of Tier one is ensured through ongoing professional development (Vaughn et al., 2007).

Tier one or the general curriculum, is designed to meet the needs of the majority of the school's population. Most students acquire basic reading skills through classroom instruction while others may struggle learning those skills from direct instruction. Universal screenings occur three times a year. These screenings assist in the early identification of students with reading difficulties and the re-identification of students who have previously had an intervention and responded (Vaughn et al., 2007). Both tier two and three involve more intensive intervention (Vaughn et al., 2007). Interventions are utilized in order for students to accelerate their progress and not fall further behind their peers. Students enter school with different levels of exposure to reading based on their experiences in their home environment. Previous reading exposure is a big factor in students making gains in basic reading skills (Al Otaiba et al., 2011a).

Interventions. As stated previously, not all students learn through general instruction, some require intervention in order to make reading gains. According to Vaughn et al. (2007), tier two involves usually about thirty minutes of additional, intensive, small group instruction that occurs daily. The classroom teacher, a specialized teacher, or another trained individual may deliver this intervention. Of the students who require tier two supports, only a small percentage will continue to show reading difficulties and require tier three intervention strategies. These interventions are typically short term meant to help boost students in a specific area. Essentially,

data gathered during the tier two process are used to make educational decisions to meet the individual student's needs in the classroom.

Intensive interventions. Tier three intervention strategies involve more explicit, more intensive, and specifically designed interventions to meet the students' individual needs. This intervention is usually performed by a reading specialist, special education teacher, or external interventionist (Vaughn et al., 2007). Research has shown that students with more individualized instruction outperform other students in reading (Al Otaiba et al., 2011a; Fuchs, Compton, Fuchs, Bryant, & Davis, 2007). For example, Fuchs et al. (2007) identified students who scored low on WIF assessments during a nine-week period. The readers who were low in reading were randomly assigned to a tutored group, as a tier two intervention, or to a non-tutored control group. The groups were assessed weekly for an additional nine weeks. Researchers also gave the students a battery of standardized reading measures. The tutored group outperformed the control group on the progress monitoring measure and the standardized reading assessments. This shows the potential positive impact of intensive interventions can have on reading skills.

Curriculum-Based Measures

RTI screeners and progress monitoring forms are usually a system of tools called curriculum-based measures (CBMs). CBMs have been created in a variety of school subject areas such as reading, writing, and math. These CBMs provide educators with a way in which to systematically track student progress and growth on measures that can be easily compared to one another. The most widely studied CBMs are in the area of reading.

CBMs are assessment tools used in the three-tiered approach of RTI. There are three functions of assessment that are necessary to a three tiered reading intervention as identified by Fuchs and Fuchs (2007). The first function of assessment is to screen the entire population not

just the target population. The next function involves monitoring progress to enable educators to make informed decisions regarding the responsiveness to intervention. The last function is to inform instructional planning in the classroom and to assist in providing individualized instruction in the classroom. Each function is a necessary component in making RTI an effective decision-making tool. CBMs are the tools that assist in gathering consist data on the population as a whole as well as to progress monitor those who have been struggling. If monitored correctly, CBMs can also inform instruction by providing data on the effectiveness of the instruction. Further, CBMs can assist in the development of student growth rate of students which has shown reliability in predicting later reading proficiency and have been shown to have strong criterion validity (Deno, 2003; Deno & Fuchs, 1991).

Deno and Mirkin (1977) first discussed CBMs and the usefulness of gathering data through these sources. CBMs are an indicator of overall system functioning for reading, writing, and math much in the way that a thermometer is an indicator of human health (Deno, 2003). In fact, Marston, Mirkin, and Deno (1984) compared CBMs against teacher referrals for evaluating for special education services. There were three groups of students; group one was referred using CBM data, group two was referred through the teacher referral process from the same schools as the CBM identification after having received some type of intervention prior to referral, and group three was comprised of students who were referred by teachers who believed the child needed special education placement and did not receive intervention prior to referral. All groups had similar numbers of initial students identified. Students were all assessed using the Woodcock-Johnson Tests of Cognitive Ability and Achievement in order to compare groups. The researchers found that group three, teacher referrals only, referred significantly more males (80%) than group one that relied on measurement data (66%). Teachers tended to refer boys

more often for school related problems. The interaction between referral and sex was significant, which the researchers hypothesized that decisions based on CBMs reduce bias based on sex. The group that was referred through continuous evaluation using CBMs qualified 80% of students eventually as Learning Disabled, whereas in the teacher referral groups, only one third qualified. These findings suggest that teacher referral is less sensitive than using CBMs. Polycn, Levine-Donnerstein, Perfect, and Obrzut (2014) also found that teacher referrals were reduced and were more accurate in identifying SLD after implementing a peer-mediated intervention for reading fluency.

CBMs have often faced scrutiny when used in a school setting. Most often, it is a misconception that reading fluency CBMs do not predict overall reading comprehension. Nonetheless, many studies have shown that for elementary age students, reading fluency as measured by CBMs is an overall good predictor of reading abilities including comprehension (Hamilton & Shinn, 2003). Hamilton and Shinn (2003) examined two groups of students, one of which they named word callers (students who could read fluently but according to teachers lacked comprehension), and the second group were peers who read with similar fluency (words read per minute), but who were judged by their teachers as being able to comprehend the material. The researchers used two types of reading CBMs and the passage comprehension subtest of the Woodcock Reading Mastery Test. The word callers did not read similarly out loud compared to the other group and had fewer correct words per minute. Teachers were not accurate in their predictions of either group's reading scores. The researchers concluded that judgment of reading abilities are often not accurate and that objective measures are better indicators.

Letter-sound fluency (LSF). Kindergarten is when children have more exposure to letters and their multiple sounds whereas first and second grades begin their exposure to short

texts. Accordingly, there are two types of CBMs typically used in kindergarten: Letter-Sound Fluency (LSF) and Letter Naming Fluency (LNF). Some districts prefer one over the other and some use both. LSF is a CBM that is used in kindergarten to track the number of letter-sounds that the child can identify in one minute. Often in kindergarten, an end of the year benchmark is 40 letter sounds per minute, LNF is another CBM that is often used as an early reading measure in preschool and/or kindergarten. Research has been mixed over which one is more predictive in future reading abilities (Stage, Sheppard, Davidson, & Browning, 2001). Stage, Sheppard, Davidson, and Browning (2001) found that albeit a small sample size (n = 59), this study found that LNF was a better predictor of first grade ORF than was LSF.

Speece and Ritchey (2005) reasoned that sound fluency is a part of the development of word reading skills that are necessary prior to reading words. Within this study, Speece and Ritchey identified at-risk and non-at-risk students in reading to decipher what skills the at-risk readers needed in their instruction in order to catch up to their peers. They found that LSF accounted for a small variance in the growth of ORF from January to May, but was not predictive of second grade. Speece and Ritchey also discussed that LNF is a skill that must be developed first, then LSF, then whole word reading, prior to passage knowledge. Further, Chall's (1983) study of reading development also discussed that the skill of letter naming must be properly developed prior to letter sounds or whole word reading.

Word-identification fluency (WIF). The skill that is next in terms of reading development following explicitly learning letter sounds, is the ability to read and identify sight words (Chall,1983). This CBM is often used as a tool for screening and monitoring reading progress in first grade students. Lists of sight words are created from high frequency word lists (Zumeta, Compton, & Fuchs, 2012). The student has one minute in which to read as many

words from the as they can while an administer marks errors on a score sheet. An end of the year benchmark for WIF is often considered to be 50 words per minute. WIF is found to have better predictive validity over nonsense word fluency or oral reading fluency (Fuchs, Fuchs, & Compton, 2004; Lopez, Thompson, Walker-Dalhouse, 2011). Fuchs, Fuchs, and Compton (2004) identified 151 children who were the lowest in their classes based on a rapid letter naming probe. The participants were then assessed using a variety of assessments including the Word Attack subtest of the Woodcock Reading Mastery Test Revised (WRMT-R WAT), Word Identification Subtest of the Woodcock Reading Mastery Test-Revised (WRMT-R WID), Comprehensive Reading Assessment Battery (CRAB), WIF CBM probe, and Nonsense Word Fluency (NWF) probe. The WRMT-R WAT and WRMT-R WID were given in both Fall and Spring while the CRAB was only given in the spring. The two CBM probes were administered for once a week for 7 weeks and twice a week for 13 weeks. In the 16 different comparisons that were run, WIF showed more predictive value over NWF in 10 of the 16 comparisons with respect to end of year reading abilities. This suggests that WIF is a better tool for assessing early reading development in first grade students.

Oral reading fluency (ORF). ORF is one of CBMs that has been most researched (Deno, 2003; Deno, Mirkin, Chiang, & Lowry, 1982; Shapiro, 1989). ORF is used as an indicator of overall reading ability (Fuchs, Fuchs, & Hosp, 2001). In an ORF CBM, students read a grade-level appropriate passage out loud to a trained person for a minute. The trained person then calculates the number of words the student read minus the errors. The score of the child falls into one of three categories: frustrational, instructional, or mastery. If the score is in the frustrational range, then that indicates that the child is not reading at their grade level (Hosp, Hosp, & Howell, 2007). There had been debate on which type of CBM to use in kindergarten as

the CBMs that are used should be developmentally appropriate to the grade level. If the CBM is not developmentally appropriate, the data gathered will not produce useful information and may just reflect floor and/or ceiling effects (Ritchey & Speece, 2006). Research has shown that ORF is not an accurate representation of a child's reading abilities until at least mid-way through the first grade (Lopez, Thompson, & Walker-Dalhouse, 2011). Lopez et al. (2011) showed for early readers, reading words in isolation was a better indicator of reading ability than reading words in context such as on an ORF passage. They postulated such findings could be due to a number of factors including: selective attention, limited automaticity in reading, and/or a developmental lag in reading.

Measuring growth. The crux of a RTI approach is gathering data and using appropriate methods to analyze the information. Growth is important in examining student data as it can help drive important instructional decisions and inform educators as to whether interventions are working. Measuring growth has been done in research and in the educational settings in a few different ways. Research on CBM measures often uses more sophisticated statistics in order to measure growth such as using growth curve modeling. Growth curve modeling is often used in longitudinal studies and is a broad term that describes a variety of statistical models for repeated measures designs (Curran, Obeidat, & Losardo, 2010; Silberglitt & Hintze, 2007). In educational settings, a slope formula is used to calculate growth by taking a final score minus initial score divided by the number of weeks in order to look at an individual student's growth over time. This gives educators an average growth rate of words or letters per week.

Growth rates vary depending on age/grade level of the child. Growth rates for children in lower grades are typically higher than those for upper as the children become more skilled readers and fits what is known regarding reading development (Fuchs, Fuchs, Hamlett, Walz, &

Germann, 1993; Petscher, Cummings, Biancarosa, & Fien, 2013; Silberglitt & Hintze, 2007). For the purposes of the current study, growth was calculated using a slope formula as growth was being compared in their second grade year. Previous research has shown that the average rate of growth for a second grade student is approximately 1.5 words per week (Deno, Fuchs, Marston, Shin, 2001; Fuchs et al., 1993; Silberglitt & Hintze, 2007; Tindal & Nese 2013).

Retention

There are many interventions that are used to address derailments in early literacy development; however, one that has been under scrutiny is grade retention. Typically, retention has been used as an intervention by educators for children who are behind in school. However, retention, regardless of when it occurs, is linked to students dropping out of high school. Stearns, Moller, Blau, and Potochnick (2007) studied school drop-out linked to retention and factors that could impact a student into dropping out of school including: SES level, ethnicity, parent involvement, and disciplinary problems. These factors did not adequately explain why retained students have such a high probability for future school dropout.

According to the United States Department of Education (2013) about 6% of kindergarteners repeat kindergarten. Children who are held back in early grades (K-2) held similar trajectories of reading growth compared to children held back in later grades (3-5). These findings fail to support the assumption that the earlier a child is held back, the more successful retention will be (Silberglit et al., 2006b). Yet, most students are held back in primary grades (Tingle et al., 2012; Warren & Saliba, 2012). Most academic gains that children make on their repeated year of kindergarten are not maintained after kindergarten (Mantzicopoulos & Morrison, 1992). The academic achievement of children who had been retained have been found to be consistently below that of their peers and retained children tend to show more aggression at

school (Jimerson & Ferguson, 2007; Tingle et al., 2012). Al Otaiba and colleagues (2011b) found that students with less formal education showed faster growth in reading in kindergarten on CBM measures than students who had attended preschool; however, students with steep growth in reading in kindergarten showed poorer performance than students with preschool experience in first grade. Researchers hypothesized that students who entered kindergarten less academically prepared, had more room to show growth, yet those students who had attended preschool ended with higher overall reading scores due to their previous reading exposure.

Silberglitt et al. (2006a) followed three groups starting in first grade until the students reached eighth grade. The three groups were students who were retained sometime between kindergarten and fifth grade, a matched group of promoted students, and a randomly selected control group. The groups were matched based on gender, school district, grade level, and reading performance as measure by a CBM. The data showed that retained students did not show the same growth rates in reading and actually made less progress than peers in later grades. By eighth grade, retained students read at a lower level than either the promoted or matched peer group. Griffith, Lloyd, Lane, and Tankersley (2010) examined growth rates for a group of students who had been retained and a group of matched students who had not been retained between kindergarten and eighth grade. Students were matched on eighth grade reading score, gender, ethnicity, and SES. The growth curves showed that the retained group had significantly lower reading achievement scores on the National Education Longitudinal Study (NELS) in eighth grade compared to the matched comparison group.

Summary and Conclusions

In summary, research has shown repeatedly over the years that retention is not a good intervention strategy, but it is still being used across the United States. One of the main reasons

that teachers retain students is that they do not want to promote students who are lacking skills that are necessary in the next grade. An underlying belief is that the student will understand the material better simply by being exposed to the material again and, thus, give children the developmental time they need in order to be successful, particularly, in reading. Additionally, the lack of intervention during the retained year (i.e., earlier grades may not utilize specialized interventions) may mean that those with similar skills who are not retained may have access to more intensified interventions in later grades sooner (Abbott et al., 2010). The purpose of the study is to examine growth rates of CBM measures comparing students who were retained to students who had similar reading scores but were promoted to the next grade. This study will help inform policy regarding retention and assist in looking more closely at students who score lower on CBM measures. Lastly, this study will examine some exploratory research on factors that may be a part of the decision making process for retention.

Chapter 3: Methods

Participants

Sample pool. Data were gathered from school records stored in the form of an electronic database from a school district in Southern Arizona. This school district has monitored basic skills in reading using CBMs. Longitudinal data were gathered from students who had scores in the frustrational range on LSF (<40 wpm) at the end of the Spring of the Kindergarten year during the 2009-2013 school years. Students also had to have second grade Spring ORF scores. They were then classified based on their retention status; the 'retained' students were those who were retained in either kindergarten or first grade and 'promoted' students who were promoted in kindergarten or first grade.

School. The school district has approximately 12,000 students enrolled from kindergarten to high school. The special education population is approximately 10% of the district population. Approximately 30% of the students in the district qualify and are set up for free and reduced lunch. Statewide percentage for free and reduced lunch is 58% (Arizona Department of Education, 2015). The student population is predominantly White, non-Hispanic (69%). The next largest ethnic group identifying as Hispanic (22%), and a small proportion of students identifying as African American (5%), Asian American or Pacific Islander (3%), or Native American (<1%). The English Language Learner (ELL) population is about 1% of the overall district population. The average population of ELL students is 9.2% nationwide and Arizona average is currently between 6.0-9.9% therefore the ELL population of this particular district is lower than average (United States Department of Education, 2015b).

Actual sample. The data were obtained from seven elementary schools within the district at the time of extraction. Students were documented as enrolled in the school in which they

attended kindergarten regardless of where they later moved schools. Seventy-two (40.9%) participants were reported to be Caucasian, 24 (13.6%) Hispanic/Latino, 7 Black/African American, and 6 (3.4%) Asian/Pacific Islander. Ethnicity was not reported by the school district for 67 (38%) students. The district reported that no students with ELL status were among those identified for the study. The first cohort (Cohort 1) entered kindergarten in the school year 2009-2010. The second cohort began in the year 2010-2011 (Cohort 2), a third cohort (Cohort 3) began in the 2011-2012 school year, and the fourth cohort (Cohort 4) entered kindergarten in the 2012-2013 school year. The students from Cohort 1 who had not been retained completed second grade in the 2011-2012 school year. The students who were retained completed second grade during the 2012-2013. Students from Cohort 2 who were not retained completed second grade in 2012-2013; those who were retained completed second grade in 2013-2014 school year. Cohort 3 completed second grade in 2013-2014 school year or 2014-2015, depending if they were promoted or retained. Cohort 4 completed second grade in the 2014-2015 school year or 2015-2016 if they were retained. Table 1 shows the number of students in the retained versus promoted groups for each cohort.

Table 1

Cohort numbers for retention

	Year started	Retained	Promoted
Cohort 1	2010	8	32
Cohort 2	2011	9	31
Cohort 3	2012	10	43
Cohort 4	2013	9	34
Total	-	36	140

Note. Retention could have occurred in kindergarten or 1st grade

The total number of students in the kindergarten data set started at 396 students based on the kindergarten criteria of students who scored into the frustrational range as defined by a LSF

score of <40. These students attended kindergarten in this school district from 2010-2013. Thirty-eight students were identified as being retained. Other kindergarten students who were in the frustrational range, but promoted, comprised 358 of the original group. Students' data were not included if they did not have scores on the Spring ORF during their second grade year. Students may have had missing data for a number of reasons including: moving out of the district, being absent during the testing day, or database error. Of the 396 total students identified in kindergarten, 176 had second grade Spring ORF scores; of those 36 (8 in kindergarten, 28 in 1st grade) had been retained and 142 had never been retained. This means that 20% of 176 students in a low performing reading group were retained. In the retained group, 17 were female and 19 were male. In the promoted group, 72 were female and 68 were male. Ten students (two from the retained group and eight from the promoted group) were excluded in analysis for Research Question #1 as missing second grade Fall ORF scores in second grade precluded the computation of reading growth. Figures 1 and 2 show participate selection for Research Questions 1 and 2.

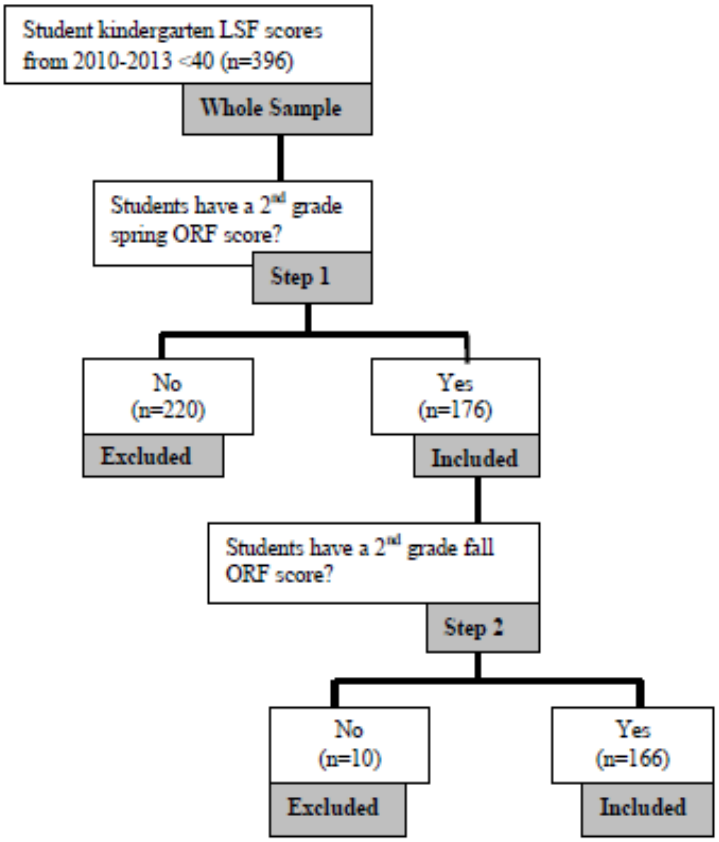


Figure 1. Flow chart of participant selection for second grade reading growth comparison

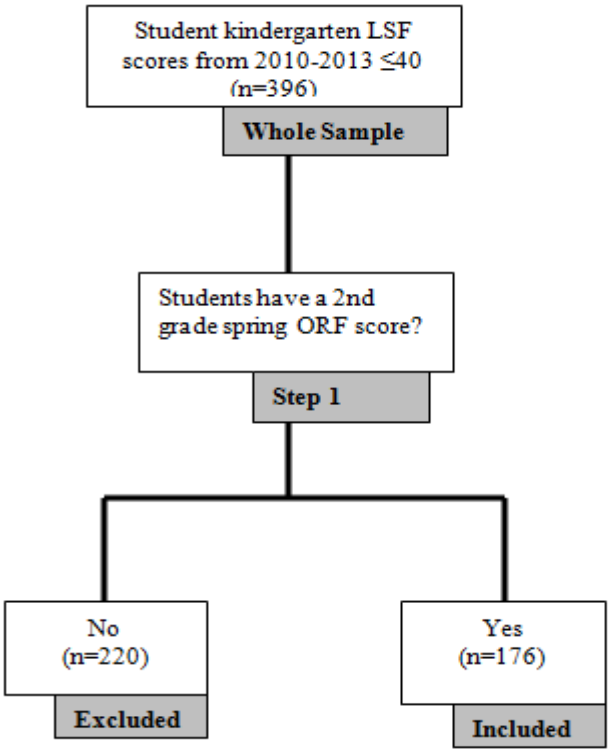


Figure 2. Flow chart of participant selection for second grade ORF score comparison between retained and promoted students

For Research Question 3, out of the 176 students, 74 students were removed from the analysis as they did not have state assessment data available. Of the 74, 64 were not available to the district (possibly moved or other unknown reason score had not been recorded) and 10 were not available as they had not yet taken the third grade reading assessment at the time of final data extraction. A subset of students (n = 102) had state assessment data available and were included in the analysis. Figure 3 shows participant selection for Research Question 3. Performance on the state assessment that was classified as "minimally proficient"/ "falls far below" would potentially result in retention in third grade as a result of the Move on Reading policy. This subset comprised the participants for Research Question #3 that examined if students who had previously been retained were still at risk of retention at the end of third grade.

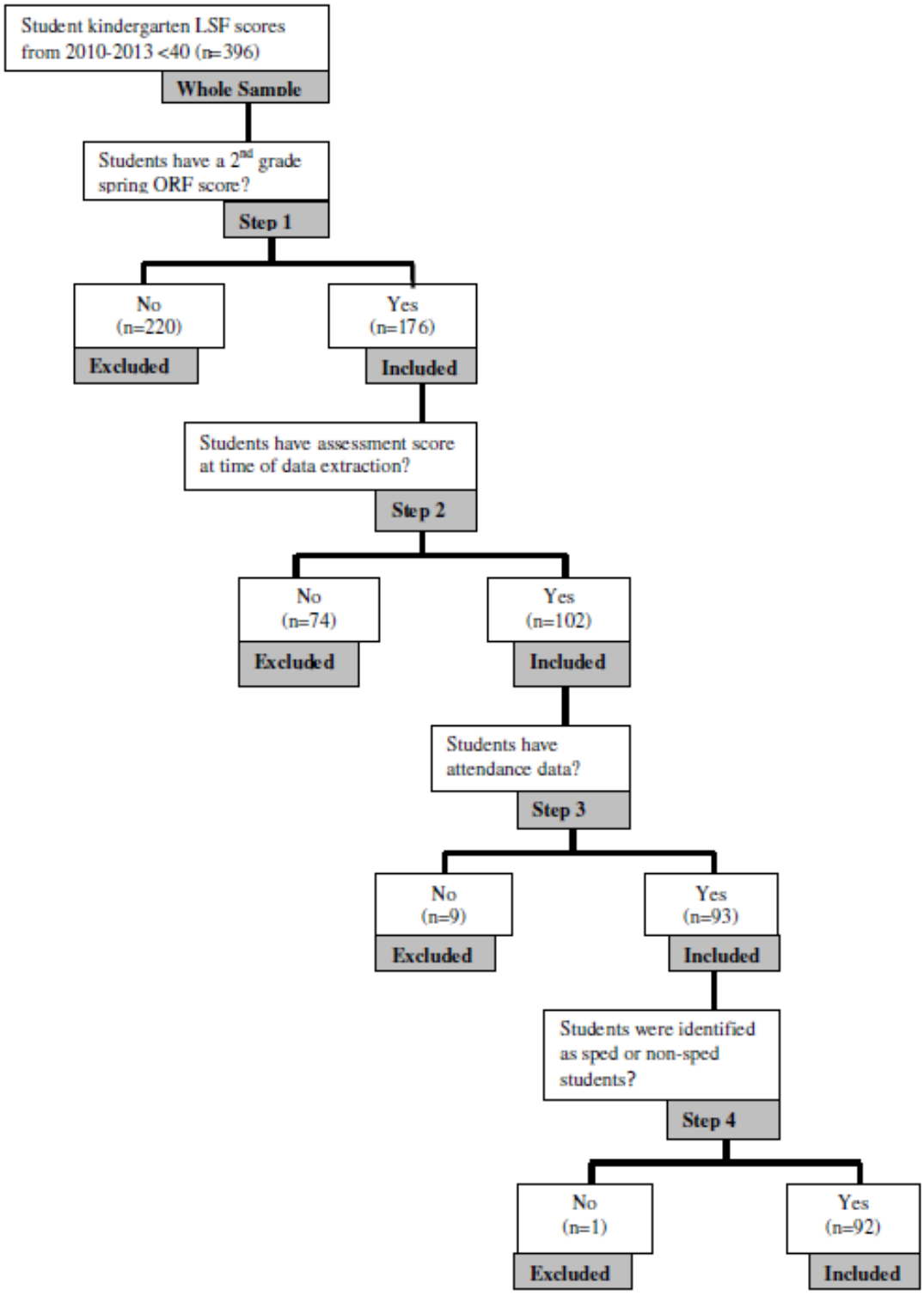


Figure 3. Flow chart of participant selection for third grade state assessment comparison

Measures

Letter-Sound Fluency (LSF). A LSF CBM was administered in kindergarten as a monthly progress monitor. Progress monitoring differs from screening in the frequency in which it is given. Progress monitoring allows for teacher responsiveness to the data that are collected and to track early reading growth. The mid-year goal of LSF is 20 letter-sounds and the end of the year goal is 40 letter-sounds. The students are shown a page with a mix of capital and lower case letters. The list that is given is consistent throughout the district. The directions state that the students must say each letter sound in order to get a point. If they do not know the letter sound, then it is provided for them. The student continues saying letter sounds until one minute has elapsed. The number of correct letter-sounds is totaled to provide them with a score. The unit of measurement is letter sounds per minute. If students are not making progress to meet the year end goal, several outcomes are possible. The teacher may change instruction, an intervention may occur, or the student may be retained. There was not a systematic process for interventions that is district wide for when students score low on LSF measures. Policies and practices differ from school to school on what is done with the data gathered from LSF. Three LSF values were recorded for descriptive purposes to characterize the sample's reading patterns: Fall LSF scores, Spring LSF scores, and LSF growth.

Word Identification Fluency (WIF). This CBM was given to first grade students district wide three times during the school year. During the assessment time, students are given one minute in which to read a list of words one-on-one with an adult trained in the assessment. The adult marks which words are read incorrectly. If a student takes more than three seconds the student is provided the word and moves to the next word. The score is calculated as correct

words per minute. The end of the year benchmark is 50 words per minute. Students' Spring WIF scores were extracted for descriptive purposes.

Oral Reading Fluency (ORF). The ORF CBM was given as a screener to students in first through fifth grade district wide, three times during the school year. A sample passage can be found in Appendix A. During this screening process, the student reads a grade level passage with a trained adult. These trained adults are teachers, reading specialists, school psychologists, and intervention specialists at the schools. The administrator follows along with the child, marking errors and supplying words when a child is stuck for at least three seconds. The score is correct words per minute (cwpm). These scores are used in a systematic way to flag students for reading interventions. Second grade students have an instructional level of 40+ cwpm. If students do not meet the instructional range, they are considered to be in the frustrational range. Students had three ORF values, ORF growth being the dependent variable for Research Question #1, and Spring ORF scores being a dependent variable for Research Question #2. Fall ORF was used to calculate growth and as a supplemental value for analysis. Growth ORF and Spring ORF scores were used as independent variables for Research Question #3.

Procedures

Historically, when the district started to implement CBMs as a school-wide assessment, the district began with ORF in first through 5th grade. Due to previous research (Lopez et al., 2011), the school district slowly implemented other CBMs to better gauge early reading. In first grade, both an ORF passage and a Word Identification Fluency (WIF) word list are used in order to more accurately gauge students' abilities in reading. WIF uses frequently used sight words on a list that students must read. LSF is implemented in kindergarten to examine early indicators of reading.

Intervention procedures. Prior to intervention, a Can't Do/Won't Do assessment is done in which the students must read the same passage, but are offered an incentive if they improve their score in order to see if motivation is the reason for the low score. If students are still in the frustrational range, then an intervention is done. The intervention for first grade is a flashcard practice with sight words. The students have three minutes to get as many words correct as they can. They are rewarded if they can beat their highest score. A generalization CBM is done every five sessions to track their overall progress. The intervention for second through fifth grades is a direct instruction reading practice. Students have a passage read to them, then they practice the passage with error correction, and then they read it independently for one minute to get a score. The students are offered prizes for each day that they beat their best score. A generalization probe or cold read is done each week in order to track their overall progress and to see if the intervention is being generalized to their overall reading on an unpracticed passage. This intervention is done for at least 15 sessions of guided practice reading with three generalizations. If students do not progress into the instructional range or show significant growth, then they are referred for special education evaluation. This has reduced the number of special education referrals in the district (VanDerHeyden, Witt & Gilbertson, 2007). The reading intervention is the same grades 2^{nd}-5^{th}. First grade is based on word lists gathered from "The Six Minute Solution: A Reading Fluency Program" by Adams and Brown (2003) as are other reading assessment passages. The school district has adapted the directions and the LSF lists to better fit the district.

Following approval from the University of Arizona Institutional Review Board and the school district, data were gathered from the System to Enhance Educational Performance (STEEP). LSF and ORF were extracted from STEEP while demographic data was gathered from

PowerSchool. Data were gathered and then another district employee de-identified the information prior to the investigator running an analysis. Placement into special education services was controlled for and analyzed for use in Research Question 3. Retention grouping was determined in two ways. The students had either second grade scores a year later than expected or had been administered the same grade level CBMs two years in a row.

Ethical and legal considerations. The Family Educational Rights and Privacy Act **(FERPA; 20 U.S.C. § 1232g; 34 CFR Part 99)** is federal law that protects the privacy of personally identifiable information within student records. FERPA requires parental permission in order to access student records. There are three exemptions to this requirement: directory information, de-identified information, and research conducted for or on the behalf of the educational institution. This study utilized a FERPA waiver under the second and the third exemptions. Another school district employee de-identified the student information and the information gathered from this study was used to examine district policies regarding retention. Students were identified with a number associated with the study and the student identification number was then deleted by a school district employee prior to analysis.

Data Analyses

For the first two research questions, a priori power analysis estimated that the minimum number of participants required was 85 when power is set to 0.80, alpha level set to 0.05, a medium effect size, and the number of potential predictors (i.e., retention status, gender, schools SES). For research question three, with number of predictors set to three, power set to 0.80, a medium effect size, and alpha set to .05, a sample size of 77 was required.

Basic descriptive statistics and reading patterns. Data regarding gender, ethnicity, English Language Learner (ELL) status, school level SES were collected for the purpose of

adding the variables into the study or to determine if variables need to be controlled in the study. Gender was included as an independent variable in the research questions and was targeted as a factor in some of the supplemental analysis to determine if differences in reading patterns existed between males and females. School level SES was included as an independent variable, coded as a percentage of students at the school with free and reduced lunch.

Additional analysis were conducted in order to better characterize reading patterns cross-sectionally and longitudinally and determine differences between the retained and promoted groups over time. Chi-square test of independence was performed to compare the frequency of retention between males and females. Chi square was also utilized for comparative differences in retention status/gender and WIF levels as well as ORF levels. An independent samples Mann Whitney was used as a nonparametric test due to unequal sizes between groups (Urdan, 2010) to compare differences between the retained and promoted groups in LSF Spring score, WIF spring score, and ORF Fall score, as these scores are continuous yet not normally distributed.

Research question 1: Reading growth. Will students who were retained demonstrate commensurate reading growth as assessed on reading CBM measures compared to promoted peers at the end of second grade? This question refers to the amount of growth obtained by students on the ORF reading measure during their second grade year. The hypothesis was that students who were retained in kindergarten or first grade would show commensurate growth on CBMs by second grade compared to students who had similar low reading scores who were promoted in kindergarten and first grade.

Analyses for Research Question 1. Data analysis for hypothesis 1 was conducted through the use of linear multiple regression as data met the following assumptions: normal distribution, linear relationship between independent and dependent variables, and

homoscedasticity (Osborne & Waters, 2002; Williams, Grajales, & Kurkiewicz, 2013). Second grade growth ratio was used as a dependent variable and measured using a slope method and expressed as a decimal. The following variables were included in the model to determine if they significantly impacted the rate of growth: sex, SES level, and retention status. Sex was coded as 0 = Female, 1 = Male, retention status was coded as 0 = promoted, 1 = retained, and social economic status measured as a percentage of students on free and reduced lunch in the school in which the child was enrolled. Second grade ORF growth ratio was calculated using a slope method. This was done by subtracting the second grade Fall ORF score from the second grade Spring ORF score and then this number was then divided by 30, which is the average number of instructional weeks between when the Fall and Spring scores were obtained.

Research Question 2: Reading Score. What are the differences in reading scores on the ORF between students who were retained in either kindergarten or first grade and their promoted peers at the end of second grade? The hypothesis was that students who were retained in kindergarten or first grade would show commensurate performance on second grade ORF compared to students who had similar low reading scores, but promoted in kindergarten and first grade.

Analyses for Question 2. To assess the relative contribution of retention grouping (retained versus not retained) in predicting second grade ORF CBM Spring scores, a multiple regression analysis was utilized using the same model as Research Question 1. Again, sex was coded as 0 = Female, 1 = Male, retention status coded as 0 = promoted, 1 = retained, and social economic status measured as a percentage of students on free and reduced lunch.

Research Question 3. Will students who were retained in kindergarten or first grade be at-risk for retention in third grade? At risk for retention will be measured by whether or not students fell into the minimally proficient (or equivalent) on the state assessment.

Analyses for Question 3. Data for the third research question was analyzed using a multinomial logistic regression (MLR) independent variable given that the state assessment score was a categorical variable with more than two groupings. This analysis was performed to model the relationship between the predictors and membership in the three groups (minimally proficient, partially proficient, and proficient). Minimally Proficient was used as the reference category. Only five students fell into the Highly Proficient category and these students were collapsed into the proficient category as the main question is pertained to the minimally proficient category as these students are at risk for retention due to the Arizona Move One When Reading Statute. The.05 criterion of statistical significance was employed for all tests. The following variables were used as independent variables in the analysis: SES level, special education status, retention status, sex, ORF Spring score, and number of absences during the course of school. Sex was coded as 0 = Female, 1=Male, retention status coded as 0 = promoted, 1 = retained, social economic status measured as a percentage of students on free and reduced lunch, and special education status coded as 0 = No, 1 = Yes. Out of the 176 original participants, 92 had complete data for this question (see Figure 3).

Chapter 4: Results

This chapter presents the statistical results of the research study. The purpose of this study was to analyze data from students who were retained and compare their reading performance to students who were promoted yet had reading scores also in the frustrational range on kindergarten LSF. The findings will be reported for each research question.

Reading Patterns of Sample

Letter-Sound Fluency. As noted in the Method section, LSF scores were used to determine eligibility for this study; that is, original Spring LSF kindergarten scores. The study's inclusionary criterion that kindergarten students had to be in the frustrational level of <40 letter sounds identified under a minute resulted in 176 students over 5 year period (refer to figures 1-3). Of those identified, 36 were retained, 8 in kindergarten and 28 in first grade. Scores ranged from 0-39 letter sounds per minute for both groups.

A Chi-square test of independence revealed that no significant difference, x^2 (1) = .203, p = .653, was observed meaning that males were just as likely to be retained as females. Results are displayed in Table 2. An independent samples Mann Whitney test indicated that in kindergarten, among students in the frustrational range, those who were retained (*Mdn* = 22) scored significantly lower on LSF than those promoted (*Mdn* = 33), U = 13.968, p < .000, two-tailed (see Table 3). The growth per week for the retained group (*Mdn* = .73) for the kindergarten LSF was also significantly less than the promoted group (*Mdn* = .92), U = 6.214, p = .013. In figures 4 and 5, the promoted group skews towards the higher end of the distribution, whereas the LSF scores of retained group are more spread out, including there being more scores towards the lower tail end of the distribution.

Table 2

Chi Square Results for Gender and Retention Status

		Gender		
		Female	Male	Total
Retained	Observed N	17	19	36
	Expected N	18.2	17.8	36
Promoted	Observed N	72	68	140
	Expected N	70.8	69.2	140
Total	Observed N	89	87	176
	Expected N	89	87	176

Note. x^2 (1) = .203, p = .653,

Table 3

Mann-Whitney Test for LSF Group Differences

	Retention status						
	Retained			Promoted			
	Mdn	Interquartile Range	n	Mdn	Interquartile Range	n	U
LSF score	22	16	36	33	10	140	13.968*

Note. LSF=end of year Letter Sound Fluency (LSF) score.

* p < .000

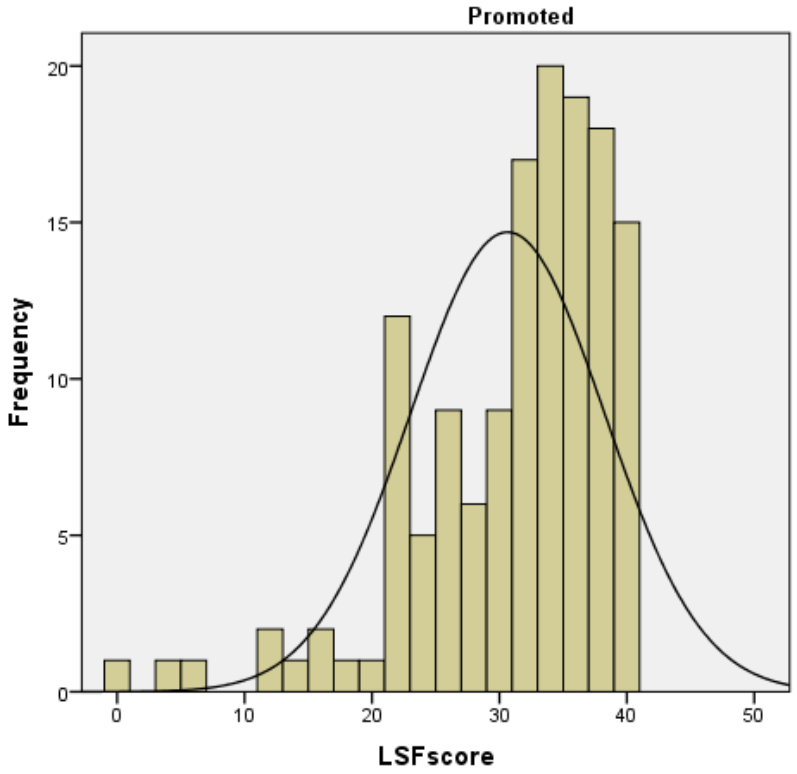


Figure 4. Letter Sound Fluency (LSF) score distribution for promoted students

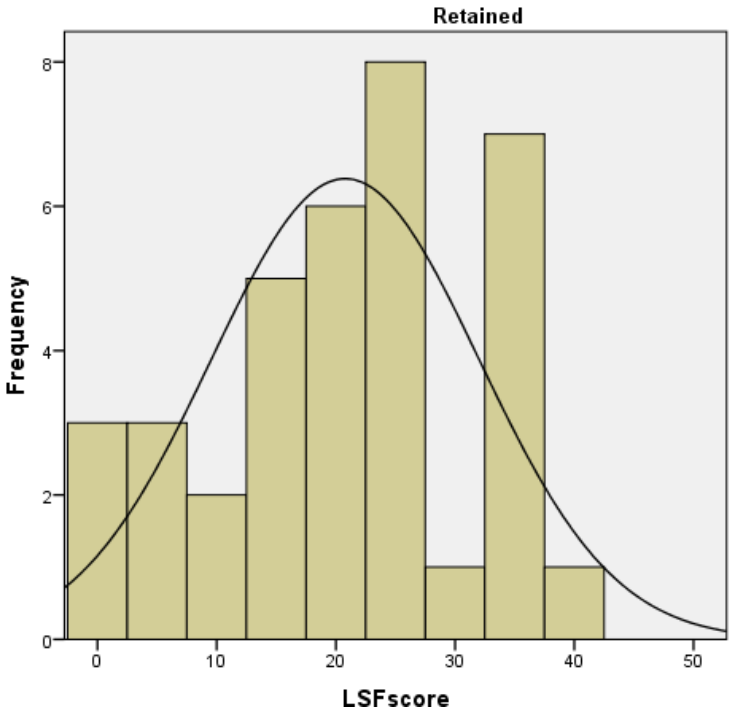


Figure 5. Letter Sound Fluency (LSF) score distribution for retained students

Word Identification Fluency. In first grade, those who had been retained (*Mdn*=32) performed significantly lower than those who had not been retained (*Mdn*=49) on their Spring WIF score and were more likely to be in the frustrational range (<50 words in a minute) than expected by chance on the Spring administration of WIF, $U = 10.520$, $p < .001$. Note that 15 students were missing a Spring WIF score. Table 4 displays these results. Both groups had a median Spring WIF score that fell into the frustrational range, with the Chi-square test of independence showing that significantly more students from the retained group were in the frustrational range on WIF than would be expected by chance, x^2 (1) = 4.052, $p = .044$ (shown in Table 5). The chi-square test of independence comparing the frequency of males and females revealed a significant interaction, x^2 (1) = 6.695, $p = .010$. Males were more likely to be in the frustrational range than females (see Table 6).

Table 4

Mann-Whitney Test for WIF Group Differences Based on Retention Status

	Retention status						
	Retained			Promoted			
	Mdn	Interquartile Range	n	Mdn	Interquatile Range	n	t
WIF score	31.87	32	31	48.63	18	131	10.520*

Note. WIF=end of year Word Identification Fluency (WIF) score.

* p < .001

Table 5

Chi Square Results for Retention Status and WIF Levels

		WIF Level		
		Frustrational	Instructional	Total
Retained	Observed N	25	11	36
	Expected N	19.6	16.4	36
Promoted	Observed N	71	69	140
	Expected N	76.4	63.6	140
Total	Observed N	96	80	176
	Expected N	96	80	176

Note. WIF=Word Identification Fluency. Frustrational level for WIF is <50 words per minute. x^2 (1) = 4.052, p = .044

Table 6

Chi Square Results for Gender and WIF Levels

		WIF Level		
		Frustrational	Instructional	Total
Female	Observed N	40	49	89
	Expected N	48.5	40.5	89
Male	Observed N	56	31	87
	Expected N	47.5	39.5	87
Total	Observed N	96	80	176
	Expected N	96	80	176

Note. WIF=Word Identification Fluency. Frustrational level for WIF is <50 words per minute. x^2 (1) = 6.695, p = .010

Second Grade ORF. Frustrational scores for second grade ORF are <40 words per minute. Fall ORF scores were analyzed to determine additional patterns. The independent samples Mann-Whitney test showed that those who were retained (*Mdn* = 33) did not score

significantly different on the Fall ORF probe than those not retained (*Mdn* = 34), *U* = .269, *p* = .604, two-tailed. Table 7 shows those results.

For Spring ORF, 23 (13.06%) out of the original 176 continued to be in the frustrational range at the end of second grade, with 7/36 (19.44%) of those who had been retained scoring in the frustrational range and 16/140 (11.42%) of those who had not been retained scoring in the frustrational range. A Chi-square test of independence was calculated comparing the frequency of retention for those in the frustrational versus instructional level on Spring ORF. No significant interaction, x^2 (1) = 1.020, *p*=.313 was observed, meaning that those who were retained or not retained were just as likely to be in the frustrational range (see Table 8). No significant interaction, x^2 (1) = .024, *p*=.877, was observed on the Chi-square test of independence comparing which gender was more likely to be in the frustrational versus instructional level on Spring ORF. Table 9 displays the results. Therefore, males and females were just as likely to be in the frustrational range.

Table 7

Mann-Whitney Test for Fall ORF Score Group Differences

	Retention status						
	Retained			Promoted			
	Mdn	Interquartile Range	n	Mdn	Interquartile Range	n	U
ORF Fall score	33	31	34	34	21	132	.141*

Note. ORF=Fall Oral Reading Fluency score.

*p = .707

Table 8

Chi Square Results for Retention Status and ORF Levels

		ORF Levels		
		Frustrational	Instructional	Total
Promoted	Observed N	18	122	140
	Expected N	19.9	120.1	140
Retained	Observed N	7	29	36
	Expected N	5.1	30.9	36
Total	Observed N	25	151	176
	Expected N	25	151	176

Note. ORF=Oral Reading Fluency. Frustrational level for ORF is <40 words per minute. x^2 (1) = 1.020, p=.313

Table 9

Chi Square Results for Gender and ORF Levels

		ORF Levels		
		Frustrational	Instructional	Total
Female	Observed N	13	76	89
	Expected N	12.6	76.4	89
Male	Observed N	12	75	87
	Expected N	12.4	74.6	87
Total	Observed N	25	151	176
	Expected N	25	151	176

Note. ORF=Oral Reading Fluency. Frustrational level for ORF is <40 words per minute x^2 (1) = .024, p = .877

Outcomes using a case-controlled approach (matched pairs). As kindergarten LSF scores were statistically different between the students who had been retained and the ones who had been promoted, further analyses were performed using case controlled match pairs (i.e., what were the outcomes in 2nd grade ORF, growth, and state standardized tests scores in 3rd grade when LSF performance between the groups were comparable. The first analysis performed was one in which retained students were paired with a promoted student first by score and had to be

within +/- 1 LS score of each other. Next, the participants were matched as closely as possible using demographic information such as sex, ethnicity, and SES level. If the retained student did not have a promoted student to be matched with based on score and at least one demographic detail, then that student was excluded from the sample. This led to 29 students in the retained group (7 excluded) and 29 in the promoted group (111 excluded). The seven from the retained group (*Mdn* = 10) did not have a promoted equivalent peer/score. The retained group had an LSF median score of 25 as did the promoted group. An independent samples Mann-Whitney was performed to compare the difference in the growth rates for the kindergarten LSF scores for the matched group. The growth per week for the retained group (*Mdn* = .80) for the kindergarten and did not significantly differ from the growth rate of the promoted group (*Mdn* = .63), *U* = 328.00, *p* = .088. For WIF, an independent samples Mann-Whitney test showed that those who were retained (*Mdn* = 36) did not score significantly different than those promoted (*Mdn* = 50), *U* = 4.464, *p* = .067, two-tailed. Further, an independent samples Mann-Whitney test for second grade ORF Fall showed that those who were retained (*Mdn* = 36) did not score significantly different than those promoted (*Mdn* = 36), *U* = .019, *p* = .895, two-tailed. A Mann-Whitney was also done comparing scores on second grade ORF between those who were retained (*Mdn* = 66) and those promoted on 2nd grade Spring ORF (*Mdn* = 72), *U* = .621, *p* = .599, two-tailed. These scores were not significant different. The last analysis using the Mann-Whitney with the matched participants was done comparing 2nd grade Spring ORF scores between the retained group (*Mdn* = .83) and the promoted group (*Mdn* = 1.08), *U* = 354.50, *p* = .692. These results were also non-significant.

Outcomes based on analysis of school. Students from the sample were from seven schools within the district. The schools all vary in population, see table 10 for percentages of

participants based on the school's population for a relative comparison. Additional analysis using independent Mann-Whitney was utilized to examine any potential differences between the seven schools in the district. No significant differences were shown for schools retained and promoted groups when examining LSF, WIF, Fall ORF and Spring ORF.

Table 10

School Proportions

	Number of Students Retained	Number of Students Promoted	Proportion
Elementary School 1	2	12	0.02
Elementary School 2	9	36	0.08
Elementary School 3	6	2	0.01
Elementary School 4	6	26	0.05
Elementary School 5	11	2	0.02
Elementary School 6	2	9	0.02
Elementary School 7	0	4	0.01

Note. Number of Students Retained and Number of Students Promoted are based on study sample. Proportion based on individual school population.

Analysis of Research Questions

Reading growth results. Were students who were retained demonstrate commensurate growth as assessed on ORF compared to promoted peers at the end of second grade? The hypothesis was that student's reading growth rates would be similar for students who were identified as low readers in kindergarten whether they were promoted or had been retained in early grades. The total number of the retained group was 34, which was comprised of 18 males and 16 females. The promoted group had a total of 132, which included 65 males and 67 females (see figure 1).

The regression was done to see if retention predicted reading growth rate [DV] when controlling for gender [IV_1] retention status [IV_2], and social economic status [IV_3]. A non-significant regression equation was found, $F(3,162) = 1.63$, $p = .185$), $R^2 = .029$. These variables in the model only explained 2.9% of the change in words read per week. Thus, being retained or not did not contribute to the model for reading growth. Table 11 shows the relative contribution of each variable entered into the model. As shown on this table, none of the factors contributed significantly to the model. The average growth rate for the promoted group was 1.04 words per week, standard deviation (SD) of .49. Among the 132 students who had been promoted, 14 students had rates of growth higher than the national growth rate of 1.5 words a week. One student had a negative growth rate as their Spring ORF score was lower than their Fall score. The average growth rate for the group of 34 students who had been retained was .93 words per week, SD of .54. Only four students had ORF growth rates above 1.5 words per week. This group had one student with negative growth and two students with zero growth.

Table 11

Reading Growth Results

	B	SE B	β
Gender	.096	.078	.096
Retention Status	-.114	.096	-.092
Social Economic Status	.708	.522	.106
R^2	0.029		
F for change in R^2	1.628		

Spring ORF scores. What are the differences in reading scores on the ORF between students who were retained in either kindergarten or first grade and their promoted peers at the end of second grade? The hypothesis predicted that no significant difference would be found in second grade ORF scores between those who had been retained versus promoted but all of whom

had been identified as struggling readers in kindergarten. This group had a total of 176 students since it was the Spring ORF that was being utilized. This group was broken down into 36 students in the retained group and 140 students in the promoted group. The 36 retained students were comprised of 17 females and 19 males. The promoted group included 72 females and 68 males (see figure 2).

Gender [IV_1], retention status [IV_2], and social economic status [IV_3] did not jointly predict Second grade Spring ORF scores (DV) when entered together in the regression equation, F (3,172) = .671, p = .571, R^2 = .012. The variables that contributed to this model only explained 1.2% of the scores on second grade ORF. No significant differences were found between students who were retained and students who were promoted in their Second grade ORF scores. Table 12 displays the results. Those who were retained (*Mdn* = 63) did not score significantly different on ORF probe than those promoted (*Mdn* = 68). Median scores show that the majority in both groups are above the instructional range of 40 words per minute.

Table 12

Differences in 2nd Grade Reading Score Results

	B	SE B	β
Gender	.949	3.344	.022
Retained	-5.749	4.140	-.106
Social Economic Status	-2.605	22.585	-.009
R^2	0.012		
F for change in R^2	0.671		

Third grade high stakes testing results. Will students who were retained in kindergarten or first grade be at-risk for retention in third grade? In order to predict whether students would be at risk for retention by falling into the minimally proficient category of the state assessment, multiple predictors (attendance, ORF second grade Spring score, school SES

level, retention status, special education status, and gender) were included into the multinomial regression model.

Table 13 displays the state test categories and how retained and promoted students fell in each category. The table shows that the five highly proficient students who were collapsed into the proficient category for statistical purposes were all in the promoted group. Thirty-six students fell into the minimally proficient group. Table 13 displays the comparison between how the research group scored on the state assessment compared to the district and then compared to the state.

Table 13

Third Grade Assessment Frequency

	Retained	Promoted	Total
Minimally Proficient	10 (10.8%)	26 (28.2%)	36 (39.1%)
Partially Proficient	4 (4.3%)	24 (26.0%)	28 (30.4%)
Proficient	8 (8.7%)	15 (16.3%)	23 (25.0%)
Highly Proficient	0 (0%)	5 (5.4%)	5 (5.4%)
Total	22 (23.9%)	70 (76.1%)	92 (100%)

Table 14

Third Grade Assessment Comparison

	Research Group[a]	District[b]	State[c]
Minimally Proficient	39%	26%	43%
Partially Proficient	31%	19%	16%
Proficient	25%	36%	30%
Highly Proficient	5%	19%	11%

[a] Research group assessment results were from state reading assessments from 2014-2017
[b] Overall district results from the 2015 state reading test results from Arizona Department of Education (2015b)
[c] Overall state results from the 2015 state reading test results from Arizona Department of Education (2015b)

As shown in Table 15, significant unique contributions were made by special education status and Spring ORF scores, $\chi 2(12, N = 92) = 82.020$, Nagelkerke $R^2 = .302$, $p = .004$. The reference group was those students who had state standardized test scores in the minimally proficient level (MP). Each predictor had two parameters, one for predicting membership in the

partially proficient (PP) group rather than minimally proficient and those who scored proficient (P). The parameter estimates are shown in Table 16. Results were significant for ORF Second grade scores showed that for every unit increase in ORF score, the score is 1.024 times more likely to fall into the Proficient category relative to the Minimally Proficient category with p=.013. Special education status was also significant and showed that student who are receiving special education services are 7.14 times more likely to fall into the minimally proficient category than student who do not receive special education services with p=.027. In examining the two predictors that are shown to be significant on the Odd Ratio (OR), when ORF second grade Spring score is higher, a student is 1.043 times more likely to score in the Proficient range than the minimally proficient range on the state assessment. The MLR was run with Fall ORF and then Spring ORF to determine if the outcome would be different. Fall ORF scores did not significantly contribute in predicting classification of performance level on the state standardized test scores. Second grade growth was also added, but Spring ORF was a more significant contributor. In regard to special education status, students were 7.14 times more likely to be in the Minimally Proficient range if they were in special education relative to those in the Proficient range.

Table 15

Predictors' Contributions in the Multinomial Logistic Regression

Predictors	X^2	df	p
Attendance	0.388	2	0.824
ORF 2nd	7.306	2	0.026*
SES	1.211	2	0.546
Retention Status	1.999	2	0.368
Sped Status	6.362	2	0.042*
Gender	0.779	2	0.677

Note. ORF 2^{nd} = 2^{nd} grade Spring Oral Reading Fluency (ORF) score. SES = Social Economic Status based on school free and reduced lunch percentages. Sped = Special education status. X^2 = amount by which -2 log likelihood increases when predictor is removed from the full model. *p<.05.

Table 16

Parameter Estimates Contrasting the Minimally Proficient Group Versus Each of the Other Groups

Predictor	Minimally Proficient vs	B	OR	p
Attendance	PP	0.021	1.022	0.598
	P	0.002	1.002	0.973
ORF Spring 2nd	PP	0.023	1.024	0.097
	P	0.042	1.043	0.013*
School SES	PP	-1.093	0.335	0.787
	P	-4.276	0.305	0.305
Retention Status	PP	0.851	2.342	0.222
	P	0.023	1.023	0.973
SPED	PP	1.112	3.04	0.102
	P	1.966	7.14	0.027*
Gender	PP	0.12	1.128	0.832
	P	0.498	1.646	0.403

Note. Retention Status scores are based on the promoted group. SPED=special education status which is based on students not receiving special education services. Gender is based on males.

Chapter 5: Discussion

This chapter provides an overview of the results and discusses how the current study contributes to existing literature. This chapter also discusses limitations of the study as well as future directions for research on retention outcomes. The negative impact of retention has a long research history (Abbott et al., 2010; Hong & Raudenbush, 2005; Hong & Yu, 2007; Jimerson, 2001; Jimerson & Ferguson, 2007; Martin, 2010; NASP, 2007; NASP, 2011; Silberglitt, Appleton, Burns, & Jimerson, 2006; Silberglitt, Jimerson, Burns, & Appleton, 2006; Stearns, Moller, Blau, & Potochnick, 2007; Tingle et al., 2012; Warren & Saliba, 2012). This study will continue to add to the research in that it is examining low readers with the retention grouping and following their reading patterns through third grade.

The current study focused on reading outcomes among students who were struggling readers during kindergarten, as determined by their performance on LSF in the frustrational range. Then, the current study examined the differences between students who had been retained versus those who had been promoted on reading performance and growth in second grade using a district-wide ORF scores. Reading scores on second grade ORF and performance on the third grade reading assessment were compared between the groups to determine whether or not retention contributed to students' reading performance. This study adds to current research by using CBM data to compare reading abilities as well as using state test scores as indicators for future risk of retention.

Overview of Findings

Most of the retained students had been retained in first grade (28 out of 36), which could be due to continued lower progression on reading skills. The first hypothesis predicted that there would not be a statistical difference between the growth rates on second grade Spring ORF for

students who had been retained in early grades compared to those who had similarly low LSF scores but who had been promoted. This question was put forth because both growth and score are important factors in examining reading performance on CBM measures (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Petscher, Cummings, Biancarosa, & Fien, 2013; Silberglitt & Hintze, 2007). The second research question examined whether or not a reading performance statistical difference would occur between the retained and promoted groups on their second grade Spring ORF. It was believed there would be no difference on ORF growth rates and Spring ORF scores because research has shown that retention does not positively impact academic performance so at the time it was thought the promoted and retained groups would have equivocal reading performance. The findings showed that as predicted, there was not a significant difference on reading growth or reading level at the end of second grade between the two groups. This indicates that for students who were retained, that retention did not lead to stronger reading skills than the promoted group. The retained group was statistically similar to the group who was low in reading but promoted to the next grade. Both groups had median scores in the instructional range. Nonetheless, supplemental analysis showed that the retained group had lower scores than the promoted group on LSF and WIF. So although the hypotheses were supported, the reasons behind may be unclear. Although it is possible that having an extra year of either kindergarten or first grade may have provided time for reading maturation to occur, the findings from Research Question 3 suggest that several students from both groups are at-risk for struggling with reading three years later, and for retention due to their test performance if not identified as special education. In this regard, when examining performance on state standardized tests as part of Research Question #3, retention did not come out as a significant predictor of state test assessment category. Rather, the two factors that contribute towards predicting which

students may fall into the Minimally Proficient category were special education status and Spring ORF scores. More specifically, however, 36 students were classified as Minimally Proficient (10 from retained group and 26 from promoted group), with 21 not classified as special education (6 from retained group and 15 from promoted group.

Reading growth rates. In regards to growth, growth rates vary depending on age/grade level of the child. Typically, growth rates are higher in beginning readers and become less noticeable as students become more skilled because advanced students have less room to grow. Children who are struggling readers, or possibly have learning difficulties, will also show less growth (Fuchs et al., 1993; Petscher et al., 2013; Silberglitt & Hintze, 2007). Both the retained and promoted groups were predicted to have no significant differences for second grade growth rates. These findings on growth fit with this research as both groups were identified as groups with struggling readers. Both the promoted and the retained groups also had means that were lower than the national average of 1.5 words per week at 1.04 and .93 words per week respectively.

Second grade Spring ORF performance. As predicted, both the retained and promoted groups were predicted to be statistically similar in their second grade Spring ORF score. This fits within current research as students who are retained often lose any of the academic benefits of being retained within 2-3 years of their retention and most retained students do not keep up (Hong & Raudenbush, 2005; Jimerson, 2001; Jimerson & Ferguson, 2007; Martin, 2010; NASP, 2007; Silberglitt, Appleton, Burns, & Jimerson, 2006; Silberglitt, Jimerson, Burns, & Appleton, 2006; Tingle et al., 2012). Given that both groups performed equally on second grade Spring ORF; it is important to note that both the LSF and WIF scores for the retained group was significantly lower than for the promoted group; however, when a matched participant analysis

was performed, no differences were apparent in either group for all data. Thus, it appears that students who perform similarly on LSF have similar trajectories regardless of whether they are promoted or retained. Such findings suggest that that retention does not lead to improved outcomes, but also is accompanied by the potential host of aforementioned negative sequelae. By second grade, retention did not strengthen reading skills beyond that of the promoted group, and the balance of gain in reading scores with other potential ramifications would need to be examined more closely.

Reading pattern analysis demonstrated that the median second grade Spring ORF scores of both groups were above the instructional range at 63 words per minute for the retained group and 68 words per minute for the promoted group. Four of 36 (11%) retained students and 10 out of 140 (7%) promoted students still fell within the frustrational range on second grade Spring ORF. Of the 14 students, 11 were identified as students receiving special education by the following year, although, it is unknown what services they were receiving in special education. For this school district, second grade ORF score cutoff for the instructional range is 40 words per minute and for third grade Fall, ORF scores classified in the instructional range increase to 70 words per minute. National norms by Hasbrouk and Tindal (2005) are widely used by reading programs. The following ORF scores are available as national norms at the 25th percentile rank: Fall-25 words correct per minute (wcpm), Winter-42 wcpm, Spring-61 wcpm. By national norms of 61 words per minute, 37 of 176 (21%) were performing at or below the 25th percentile in reading.

Research Question 3. This question explored whether children who had been previously retained would be at risk of retention again in third grade due to a failure to meet the state test requirement for promotion. Students with scores classified as Minimally Proficient on the state

assessment are at risk of retention due to the Arizona Move on When Reading statute. This statute does have exemptions such as if the child has been diagnosed with a significant reading impairment or has a disability and the IEP team agrees on promotion. Approximately 70% (64/92) of students who had been identified as low readers in kindergarten who had state assessment scores, fell into either the Minimally Proficient or Partially Proficient range on the state assessment meaning that these students continued to struggle whether or not they had been previously retained. Thirty-six of the 92 students scored into the Minimally Proficient range on the state assessment meaning that they would be at risk for retention. Of those 10 had been retained; 4 of those who had been retained had been deemed eligible for special education, which is one of the exemptions for mandated retention. However, six (6.5% of the 92) students in the retained group were not identified as special education students, meaning that they could potentially be retained again. None of these six students were in the frustrational range on their Spring ORF measure.

The analysis of this question found that the model with the independent variables of retention status, gender, SES status, attendance, Spring ORF score, and special education status significantly predict for classification on the state reading test. In particular, the factors of special education status and Spring second grade ORF score showed significant predictive power. The variable of special education status showed that students who were identified as receiving special education services did worse on the state assessment than those who were not receiving services. This finding makes sense as students who have been identified as needing special education services are often provided those services due to difficulties in reading. Students with disabilities often score lower than students without disabilities on standardized assessments such as the National Assessment of Educational Progress (NAEP). Further, the NAEP has shown this pattern

for students with disabilities on the last three administrations of this assessment (National Center for Learning Disabilities, 2014).

Administering the Spring ORF has strong predictive power in identifying students at-risk. The question then becomes are lower reading scores due to a learning disability or some other factor that could be addressed through intervention before taking the third grade state standardized reading assessment. Previous research has shown moderate to strong correlations between the CBMs and state standardized testing. The strongest correlation (approaching .70) was between ORF and the state assessment; although, it was noted that hierarchical regression showed that the Winter CBM was the strongest predictor (Shapiro et al., 2006). The significance of the current study's finding that ORF Spring score showing predictive power is that this can assist with identifying students who may need a more intensive intervention leading up to the third grade state assessment in order to prevent them from being at risk for retention.

Significance of reading patterns and supplemental analyses. Additional analysis were performed in order to better understand the nature of differences between the retained and non-retained groups and factors that contribute to continued reading performance difficulties. Accordingly, there were group differences noted on the LSF probe that initially identified students as participants for this study. The group that was retained was significantly lower than the promoted group as the mean was 10 letter sounds lower for the retained group. The promoted group had a median score of 30 compared to the retained group with a median of 20 letter sounds.

As the groups progressed to first grade, there continued to be significant differences between the two groups. Just as in kindergarten, those who were in the retained group continued to have a significantly lower score on that grade level CBM WIF. Further, they also

demonstrated significantly lower growth rates. Consequently, students who were retained had evidenced poorer reading skills than those who were promoted. By second grade, there were no longer significant differences between the two groups in regards to Spring ORF reading scores or reading growth. Importantly, the findings indicate that the majority (86%) of all students who had been identified as struggling readers in kindergarten scored in the instructional range at the end of second grade. This may be misleading as in the subsequent year, performance on the state standardized test indicates students in both groups are struggling readers.

The apparently at-level reading performance of the majority of the students could potentially indicate several possibilities. One is that both groups continue to struggle in reading despite scoring into the instructional range as cutoff scores can differ. As stated previously, when compared to national norms, 21% of students were reading lower than the 25th percentile. Consideration that these students no longer exhibited reading difficulties should be done so cautiously as it important to keep in mind that cutoffs for instructional can vary depending on the district and the percentile cutoff that they have chosen. School districts may choose a particular cutoff depending on the purpose of the CBM measures and whether they are used for identification of disabilities or for identification of students who need reading interventions. As previously mentioned, another reason for these fluctuations is due to the way in which participants were included in this study. When students were matched by score and demographics, there were no longer any significant differences between the groups in LSF, WIF, Fall ORF, or Spring ORF.

Continued struggles with reading. When examining the second grade results, as previously noted, 23 (13.06%) out of the original 176 continued to be in the frustrational range (<40 wpm) in Spring, with 7/36 (19.44%) of those who had been retained scoring in the

frustrational range and 16/140 (11.42%) of those who had not been retained scoring in the frustrational range. Eleven out of these 23 students scored into the Minimally Proficient category on their state standardized assessment and all but one were identified as receiving special education services. Of the 36 students in the Minimally Proficient category on the state test, 18 were not identified as receiving special education services. Some of those 18 students, who were in the Minimally Proficient category on the state reading test, may have been struggling students who require specialized instruction, but that conclusion leaves out many factors which can impede performance. Other factors are not know as to whether students were in the process of being identified for special education, whether parents had previously refused services, or if students had exclusionary factors that prevented services from being provided. The main implication from this study is that students who were retained did not have significantly different outcomes than students who were promoted that had equally low kindergarten reading, which calls into question whether the potential negative impacts of retention outweighed other potential interventions that may have had a more positive outcome.

Researchers have postulated that neither grade retention nor promoting students positively impact a child's overall academic progress as neither promotion nor retention address the child's areas of difficulty. Researchers have concluded that the specific interventions that are utilized to advance the child's academic or behavioral skills may hold a lasting and more powerful impact (NASP, 2007). Targeting specific difficulties in order to develop a plan when a student is struggling in reading often has less negative effects according to research on retention. The system of Response to Intervention utilizes evidence based interventions and systematically tracks data which has been demonstrated as a successful method of targeting those specific deficits. The current study does not clearly show positive or negative impacts from retention as

neither group showed significant gains; therefore, the continued interventions focused on the area of deficit would be the best alternative to retention.

Limitations of the Present Study

Several limitations have been identified in the present study that suggests that caution should be taken when interpreting the data.

Sampling. The first limitation identified is that all students in the study were from one school district in Arizona that had an existing database. Working with an already existing database has many limitations as there may be information that is difficult to obtain from the sample either because it did not exist in the existing database or information had to be compiled from several different databases. Participants were excluded due to missing portions of data. As this was a study looking over students' data over several years, many students had moved or were missing data that were considered vital to the present research. Furthermore, students all received the same interventions without individualization. The students' progress through the interventions is not kept within the database so it is unknown whether or not the students made adequate progress within the intervention. Special education status was not consistently available, but this was manually extracted for students who had state assessment scores. Further, this study did not track the category of special education status for the students or what services they received. This study could have participants included who are classified as special education students but do not receive any support for reading.

Several variables were not included in this study as factors that could have impacted the findings. One was that it was also not possible to obtain individual student SES levels through the existing database. The only information of this nature that could be gathered was school level SES status using free and reduced lunch status. Students who receive free and reduced lunch

score lower on reading assessments (U.S. Department of Education, 2011). Knowing individual free and reduced lunch status would have added another factor to the present study. The SES levels appeared homogenous. Although certain schools may contain higher SES students, much of the individual student information is lost as it is unknown whether in this study home SES status impacted student's reading. Students who enter kindergarten with lower reading abilities tend to have that reading gap widen over time. That disadvantage could be environmental factors (McLoyd & Purtell, 2008). SES factors such as parent's level of education and employment status have been linked to an increased likelihood for retention (Gonzalez-Betancor & Lopez-Pulg, 2016). Family income has also been linked to outcomes on standardized reading assessments (Annie E. Casey Foundation, 2009). Too many students had missing information in this area of ethnicity so that piece of demographic information was not included as a factor. Data has shown that those who identify as Black, Hispanic, or American Indian show trends of scoring lower on standardized reading assessments than their White peers. These minority groups score below proficient at levels of around 80% whereas those who identify as White score below proficient at 58% (Annie E. Casey Foundation, 2009). Groups for the study, in particular research question 3, were limited. This was due to whether students had aged into the grade in which data were being pulled at that time.

Measurement. Several assumptions were made when choosing the groups for this study. As reading is a big learning component in early school years, it was assumed that low performance in reading was a factor in the decision to retain students. It is unknown whether reading was one of the determining factors for retention or whether behavior had an impact in the schools' decision. Retentions also can be requested by parents and not initiated by the school itself. Retention also occurs for reasons other than reading, so by focusing on only low readers

identified by their low kindergarten CBM, this study could have missed other groups of students who were retained.

This district has a majority of White non Hispanic students and a low ELL population in comparison to other school districts in southern Arizona This lack of diversity within the district may mean that the results from this study may not be applicable to all groups as ELL students may do more poorly on measures of reading depending on their level of English acquisition. Although it may be noted that ELL learners who become poor readers have difficulties with aspects of language (Geva & Massey-Garrison, 2012). Studies on retention have also shown minorities and ELL students are far more likely to be retained than their peers (Anderson et al., 2002; Tingle et al., 2012).

Future Directions

Overall, there may be sociocultural factors that contribute to the decision to retain as well as the effects (benefits, harms, or neutral impact) of retention. The district from the current study uses Light's Retention Scale which does include questions in some of these areas; however, as this scale uses a total amount of points from all areas, it may not eliminate sociocultural factors completely. These factors could be ELL status, multiple metrics of SES levels (e.g., parent education level, employment status, and household income), as these factors have been found to predict early reading skills (Eklund et al., 2013; Manolitsis et al, 2013: Massey-Garrison, 2012). This is imperative as early reading skills can predict future reading abilities (Cunningham and Stanovich, 1997).

Future studies should continue to investigate the impact of early reading and early retention on these high stakes assessments. Cunningham and Stanovich (1997) found that early reading difficulties contributed to academic issues through 11th grade. Retention has also been

shown to have later negative impacts on education and self concept (Abbott et al., 2010; Anderson et al., 2002; Goos et al., 2013; Hong & Raudenbush, 2005; Hong & Yu, 2007; Jimerson, 2001; Jimerson & Ferguson, 2007; Martin, 2010; NASP, 2007; Silberglitt, Appleton, Burns, & Jimerson, 2006; Silberglitt, Jimerson, Burns, & Appleton, 2006; Stearns, Moller, Blau, & Potochnick, 2007; Tingle et al., 2012). Early intervention in reading rather than intervention through retention can potentially assist children in the underlying cause of their reading difficulties and prevent retention via the third grade state assessment. Changes in state assessments should be monitored through this lens to provide additional insight into the reliability of these high stakes assessments.

Additionally, longitudinal studies should continue to occur following lower achieving students and those who have been retained to see further out effects regarding state testing as studies have found that most often any academic benefits from retention fade over time (NASP 2011; Wu, West, & Hughes, 2007). In order to better investigate retention, students should be included who are reading at grade level and who are strong readers. Studies can investigate whether retained students continue to underperform each year in the state assessment compared to students who were promoted. This can also inform educators not only about retention but about changes in curriculum. Curriculum can impact the way in which early reading skills are taught and continued research needs to continue in order to study the effectiveness of that curriculum.

The focus of this study was not only on retention, but also to examine why some students are who are poor readers are retained and others are not. There may be moderator variables that interact with retention and influence subsequent reading skills. Parental factors were not able to be addressed in this study. These factors have been shown to significantly impact whether or not

a student is retained. Children are more likely to be retained if parents have a lower IQ, if parents are less involved in school, and increases with number of moves a child has made (Jimerson & Kaufman, 2003; McCoy & Reynolds, 1999). The reasons for retention are often not documented through data bases. Some districts may document reasons in the students' permanent file that make it difficult to examine patterns and trends with retention. Additionally, districts should follow patterns in retention as this practice continues to show trends that more boys, ELL students, minorities, children with learning disabilities, and children with early birthdays continue to have much higher rates of retention (Gonzalez-Betancor & Lopez-Pulg, 2016; Tingle et al., 2012).

Conclusion

In summary, retention has been shown to have a consistently negative impact on students and may not target underlying needs the student has. This study compared students who had performed similarly on early CBM measures who had either been retained or promoted. This study showed that no significant differences were present in these two groups when scores were compared two years later on CBMs. When state testing scores were also compared, some discernible differences were seen. Little research has been done to date regarding whether or not these students would be potential at risk for a second retention. The current study did not show a statistical difference with the added data of the state assessment scores; however, it showed that a high percentage of students who had lower reading scores in early grades were highly likely to score below the proficient level on the state assessment. Critically, in order to be considered a successful intervention, students who were retained not only needed to exceed the performance of those who had been promoted, but also had to not still exhibit reading challenges. Even if they performed worse at study initiation and then performed comparable too to their promoted

classmates, if they still had difficulty with reading, then data would support both groups warranted further intervention or consideration for special education eligibility and services. Thus, findings indicated that retention did not benefit that group of students as their scores were statistically the same as the group that was promoted. Retention did not significantly increase reading scores and many students in this study also scored into the Minimally Proficient category in their 3rd grade reading assessment meaning that they could potentially be at risk for retention again. Continued research in this area still needs to occur as retention is still being used as an intervention in many different states across the country.

Appendix A

Example of Letter Sound Fluency Script

Scripted Instructions for Letter Sound Fluency CBM Probe Administration

Provide the class with an independent seatwork assignment. Call individual students away from the group to read with the teacher one on one.

Place the copy of the student sheet in front of the student and the examiner copy on a clipboard out of student sight.

Write the student's name and grade, teacher name, and today's date on the top of the examiner copy.

Set the countdown timer for 1 minute.

Say, **"Here are some letters** (point to the student copy). **Begin here** (point to the first letter) **and tell me the SOUNDS of as many letters as you can. If you come to a letter you don't know, I'll tell it to you. Any questions? Put your finger under the first letter. Ready? Go."**

Follow along on the examiner copy as the student says the letter sounds and mark a slash (/) through any incorrect letters. When the timer beeps at the end of 1 minute, put a bracket after the last sound provided. Thank the student, then score the probe.

Special Administration and Scoring Considerations for LSF CBM:

1. If the student says the letter name instead of the letter sound, say: **"Remember to tell me the sound the letter makes, not its name."** This is provided one time only.
2. If the student makes an error, he is not corrected. The only time a letter sound is provided is if the student hesitates for 3 seconds.
3. If the student skips a row, draw a line through it and do NOT count it in the scoring as attempted or as errors.
4. The capital letter I and lowercase L look alike, so either sound is considered correct.
5. If the student does not get any sounds correct in the first row, discontinue the task.
6. If the student finished in less than 1 minute, note the number of seconds it took to complete the sheet and prorate the score using this formula:
    a. Total number of sounds read correctly divided by the number of seconds multiplied by 60 = estimated number of letter sounds in 1 minute (e.g., 20 sounds divided by 50 seconds times 60 = 24 est. letter sounds in 1 minute)

Appendix B

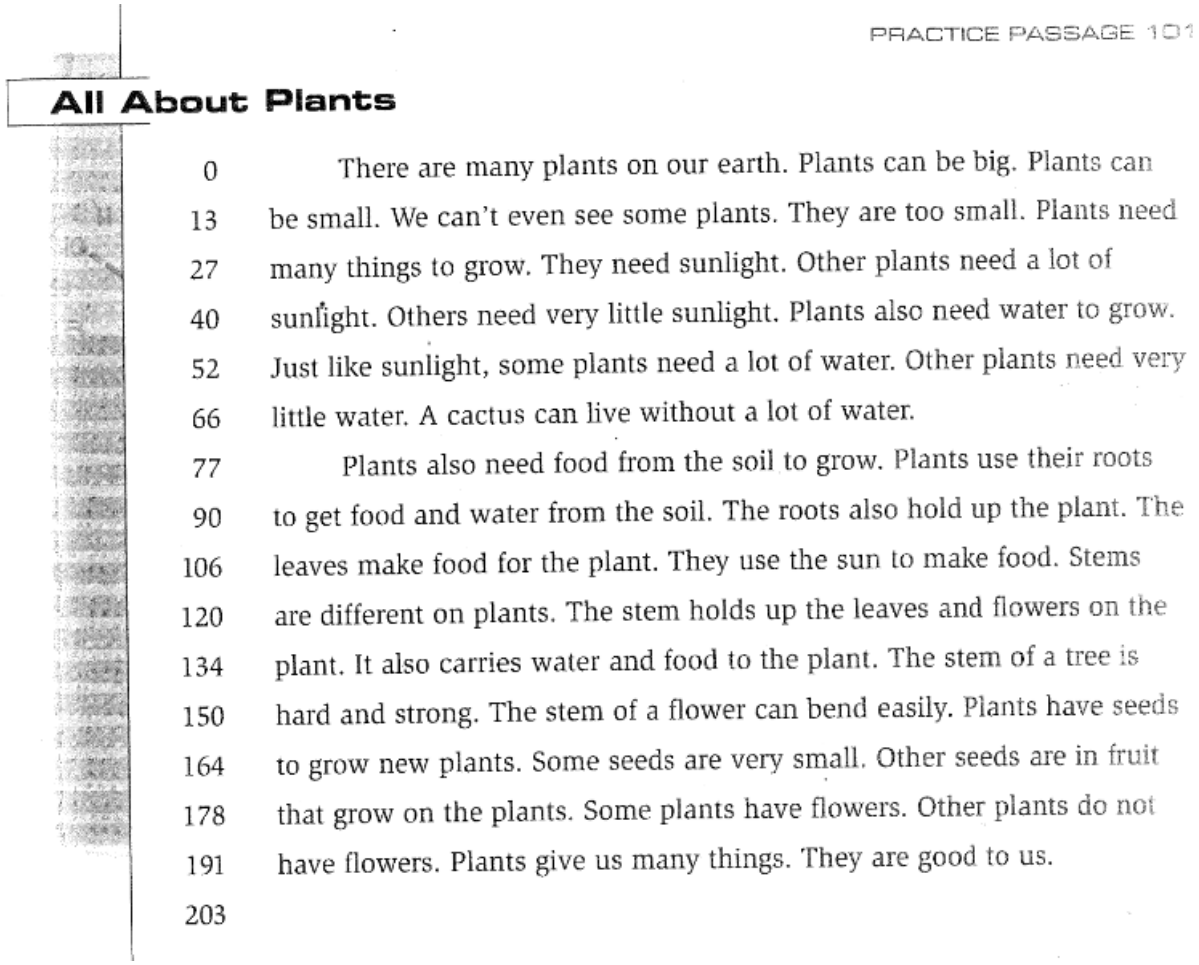
Example Letter Sound Fluency Probe

List 6

C	F	W	M	q	I	r	G	w	t
X	j	K	l	z	H	q	S	x	J
B	S	f	z	F	R	V	g	D	h
Z	x	C	Y	b	v	M	n	d	O
c	W	v	P	n	B	Y	m	U	f
H	e	R	V	y	Q	u	X	k	t
I	L	o	p	D	k	I	a	E	s
T	A	e	K	f	N	g	J	d	l
Q	c	u	N	P	O	y	A	p	i
T	G	r	E	a	s	W	d	L	T

Appendix C

Example ORF Passage



PRACTICE PASSAGE 101

All About Plants

There are many plants on our earth. Plants can be big. Plants can be small. We can't even see some plants. They are too small. Plants need many things to grow. They need sunlight. Other plants need a lot of sunlight. Others need very little sunlight. Plants also need water to grow. Just like sunlight, some plants need a lot of water. Other plants need very little water. A cactus can live without a lot of water.

Plants also need food from the soil to grow. Plants use their roots to get food and water from the soil. The roots also hold up the plant. The leaves make food for the plant. They use the sun to make food. Stems are different on plants. The stem holds up the leaves and flowers on the plant. It also carries water and food to the plant. The stem of a tree is hard and strong. The stem of a flower can bend easily. Plants have seeds to grow new plants. Some seeds are very small. Other seeds are in fruit that grow on the plants. Some plants have flowers. Other plants do not have flowers. Plants give us many things. They are good to us.

References

Abbott, M., Wills, H., Greenwood, C. R., Kamps, D., Heitzman-Powell, L., & Selig, J. (2010). The combined effects of grade retention and targeted small-group intervention on students' literacy outcomes. Reading & Writing Quarterly, 26, 4-25. Doi: 10.1080/10573560903396876

Adams, G. N. & Brown, S. M. (2003). The six-minute solution: A reading fluency program. Longmont, CO: Sopris West.

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2011a). Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. *The Elementary School Journal, 111*, 535-560.

Al Otaiba, S., Folsom, J. S., Schatschneider, C., Wanzek, J., Greulich, L., Meadows, J., Li, Z., & Connor, C. M. (2011b). Predicting first-grade reading performance from kindergarten response to tier 1 instruction. *Council for Exceptional Children, 77*, 453-470.

Annie E. Casey Foundation. (2010). Early warning! Why Reading by the end of third-grade matters: A kids count special report from the Annie E. Casey Foundation. Baltimore, MD: Annie E. Casey Foundation.

Annie E. Casey Foundation. (2012). Double jeopardy: How third-grade reading skills and poverty influence high school graduation. Baltimore MD: Donald J. Hernandez.

Arizona Department of Education. (2015a). Percentage of children approved for free or reduced-price lunches [table]. Retrieved from: http://www.azed.gov/health-nutrition/files/2012/07/sy2015_frpercentages_publish2015.8.23.pdf

Arizona Department of Education. (2015b). Retrieved from: https://www.azed.gov/accountability-research/data/

Beebe-Frankenberger, M., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Sorting second-grade students: Differentiating those retained from those promoted. *Journal of Educational Psychology, 96,* 204-215.

Brown-Chidsey, R. & Steege, M. W. (2010).*Response to intervention: Principles and strategies for effective practice, 2nd edition*. New York, NY: Guilford Press.

Chall, J.S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.

Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A., … Crouch, R. C. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology, 102,* 327-340.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98,* 394-409.

Cunningham, A. E. & Stanovich K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934-945.doi: 10.1037/0012-1649.33.6.934

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognitive Development, 11*, 121-136. doi:10.1080/15248371003699969

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232.

Deno, S. L. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention, 28*, 3-12.doi: 10.1177/073724770302800302

Deno, S. L. & Fuchs, L. S. (1991). The effects of curriculum within curriculum-based measurement. *Exceptional Children, 58*, 232-243.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*, 507-524.

Deno, S. L., Mirkin, P. K. (1977). *Data-based program modification: A manual*. Council. Reston, Virginia: Council for Exceptional Children.

Deno, S. L., Mirkin, P. K., Chiang, B, & Lowry, L. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-45.

Education Commission of the States. (2014). *Third-grade reading policies*. Denver, CO: Emily Workman.

Education Commission of the States (2016). *K-3 Quality: Does the state require full-day kindergarten?* Denver, CO: Retrieved from: http://ecs.force.com/mbdata/mbquestRT?rep=KQ302.

Eklund, K. M., Torppa, M., & Lyytinen, H. (2013). Predicting reading disability: Early cognitive risk and protective factors. *Dyslexia, 19*, 1-10.

Fletcher, J.M., Coutler, W.A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54*, 304-331.

Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1992). The validity of discrepancy-based definitions of reading disabilities.*Journal of Learning Disabilities, 25*, 555-561.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., &Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of learning Disabilities, 38*, 98-108.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59-109.

Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, J., & Davis, G. N. (2008). Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing, 21*, 413-436.

Fuchs, D. & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice, 22*, 129-136.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93-99.

Fuchs, L. S., & Fuchs, D. (2007). The role of assessment in the three-tiered approach to reading instruction. In D. Haager, J. Klinger, & S. Vaughn (Eds.), *Evidence-based reading practices for response to intervention* (3-10). Baltimore, Maryland: Paul H. Brookes Publishing Co.

Fuchs, L.S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71*, 7-21.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect*? School Psychology Review, 22,* 27-48.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239-256.

Fulmer, S. M., & Tulis, M. (2013). Changes in interest and affect during a difficult reading task: Relationships with perceived difficulty and reading fluency. *Learning and Instruction, 27, 11-20.*

Geva, E. & Massey-Garrison, A. (2012). A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of Learning Disabilities, 46*, 387-401. DOI: 10.1177/0022219412466651

Gonzalez-Betancor, S. M., & Lopez-Pulg, A. J. (2016). Grade retention in primary education is associated with quarter birth and socioeconomic status. *PLOS ONE, 11*. DOI: 10.1371/journal.pone.0166431.s001

Goos, M., Schreier, B. M., Knipprath, H. M. E., De Fraine, B., Van Damme, J., & Trautwein, U. (2013). How can cross-country differences in the practice of grade retention be explained? A closer look at national educational policy factors. *Comparative Education Review, 57*, 54-84.

Griffith, C. A., Lloyd, J. W., Lane, K. L., &Tankersley, M. (2010). Grade retention of students during grades k-8 predicts reading achievement and progress during secondary schooling. Reading & Writing Quarterly, 26, 51-66. DOI: 10.1080/10573560903396967

Hamilton, C. & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. School Psychology Review, 32, 228-240.

Harris, J., Botting, N., Myers, L., & Dodd, B. (2011). The relationship between speech impairment, phonological awareness and early literacy development. *Australian Journal of Learning Difficulties, 16,* 111-125.DOI: 10.1080/19404158.2010.515379

Hasbrouck, J. & Tindal, G. (2005) *Oral reading fluency: 90 years of measurement*. Eugene: Behavioral Research & Teaching.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hong, G. & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27,* 205-224.

Hong, G. & Yu, B. (2007). Early-grade retention and children's reading and math learning in elementary years. *Educational Evaluation and Policy Analysis, 29,* 239-261. DOI: 10.3102/0162373707309073

Hosp, M., Hosp, J., & Howell, K. (2007).*The ABCs of CBM: A practical guide to curriculum-based measurement*. New York, NY: Guilford Press.

Individuals with Disabilities Education Act (2006). Washington, D. C.: Department of Education. Retrieved from http://idea.ed.gov

Indrisano, R. & Chall, J. S. (1995). Literacy Development. *Journal of Education, 177,* 64-83.

Jimerson, S.R. (2001). Meta-anaylsis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30,* 420-437.

Jimerson, S.R., Anderson, G., & Whipple, A. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools, 39,* 441-457.

Jimerson, S. R., & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School Psychology Quarterly, 22*, 314-339.DOI: 10.1037/1045-3830.22.3.314

Jimerson, S.R. & Kaufman A.M. (2003). Reading, writing, and retention: A primer on grade retention research. *The Reading Teacher, 56,* 622-635.

Jimerson, S. R., Pletcher, M. W., Graydon, K., Schnurr, B. L., Nickerson, A. B., & Kundert, D. K. (2006). Beyond grade retention and social promotion: Promoting the social and academic competence of students. *Psychology in the Schools, 43,* 85-97. DOI: 10.1002/pits.20132.

Judge, S. & Bell, S. M. (2011). Reading achievement trajectories for students with learning disabilities during the elementary school years. *Reading & Writing Quarterly, 27,* 153-178. DOI: 10.1080/10573569.2011.532722

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437-447.

Keller-Margulis, M., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review, 37,* 374-390.

Leontopoulou, S., Jimerson, S.R., & Anderson, G.E. (2011). An international exploratory investigation of students' perceptions of stressful life events: Results from Greece,

Cyprus, and the United States. *School Psychology International, 32*, 632-644. DOI: 10.1177/0143034311403059

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36*, 596-613. DOI: 10.1037//0012-1649.36.5.596

Lopez, F. A., Thompson, S. S., & Walk-Dalhouse, D. (2011). Examining the trajectory of defferentially skilled first graders' reading fluency of words in isolation and in context. *Reading & Writing Quarterly, 27*, 281-305. Doi: 10.1080/10573569.2011.596095

Madelaine, A., & Wheldall, K. (2005). Identifying low-progress readers: Comparing teacher judgment with a curriculum-based measurement procedure. *International Journal of Disability, Development, and Education, 52*, 33-42.

Manolitsis, G., Georgiou, G. K., & Tziraki, N. (2013). Examining the effects of home literacy and numeracy environment on early reading and math acquisition. *Early Childhood Research Quarterly, 28*, 692-703.

Mantzicopoulos P. & Morrison D. (1992). Kindergarten retention: Academic and behavioral outcomes through the end of second grade. *American Educational Research Journal, 29*, 182-198.

Martin, A. J. (2010). Holding back and holding behind: Grade retention and students' non-academic and academic outcomes. *British Educational Research Journal, 5*, 739-763.

Mathes, P.G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., &Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*, 148-182. DOI: 10.1598/RRQ.40.2.2

McCoy, A. R. & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology*, *37*, 273-298.

McLoyd, V. C., & Purtell, K. (2008). How childhood poverty and income affect children's cognitive functioning and school achievement. In S. Neuman (Ed.), Educating the other America: Top experts tackle poverty, literacy, and achievement in our schools. (pp.53-72) Baltimore: Paul H. Brookes.

Morris D., Shaw, B., & Perney, J. (1990). Helping low readers in grades 2 and 3: An after-school volunteer tutoring program. *The Elementary School Journal, 91*, 132-150.

National Assessment of Educational Progress. (2013). [Graph illustration] *Percentage of students at or above Proficient in fourth- and eighth-grade NAEP mathematics and reading: various years, 1990–2013.* Retrieved from http://nationsreportcard.gov /reading_math_2013/#/what-knowledge

National Association of School Psychologists (2007). Grade retention and promotion: Information for parents. Bethesda, MD: National Association of School Psychologists.

National Association of School Psychologists (2011). NASP Position Statement: Grade retention and social promotion. Bethesda, MD: National Association of School Psychologists.

National Center for Learning Disabilities. (2014). *The state of learning disabilities: Facts, trends and emerging issues (3rd edition).* New York: National Center for Learning Disabilities Inc.

Ortiz, M., Folsom, J. S., Al Otaiba, S., Greulich, L., Thomas-Tate, S., & Connor, C. M. (2012). The componential model of reading: Predicting first grade reading performance of culturally diverse students from ecological, psychological, and cognitive factors assessed

at kindergarten entry. *Journal of Learning Disabilities, 45*, 406-417.DOI: 10.1177/0022219411431242

Osborne, J. W. & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation, 8*. Retrieved from: http://pareonline.net/getvn.asp?n=2&v=8.

Petscher, Y., Cummings, K. D., Biancarosa, G., & Fien, H, (2013). Advanced (measurement) qpplications of curriculum-based measurement in reading. *Assessment for Effective Intervention, 38*, 71-75. DOI: 10.1177/1534508412461434

Polcyn, D. M., Levine-Donnerstein, D., Perfect, M., & Obrzut, J. E. (2014). Reading intervention and special education referrals. *School Psychology Forum, 3*, 156-167.

President's Commission on Excellence in Special Education. (2002). A new era: Revitalizing special education for children and their families. Washington, DC: U.S. Department of Education.

Range, B. G., Holt, C. R., Pijanowski, J., & Young, S. (2012). The perceptions of primary grade teachers and elementary principals about the effectiveness of grade-level retention. *The Professional Educator, 36*.

Rasinsky, T. (2004). Creating Fluent Readers. *Educational Leadership, 61*, 46-51.

Ritchey, K. D., & Speece, D. L. (2006). From letter names to word reading: The nascent role of sublexical fluency. *Contemporary Educational Psychology, 31*, 301-327.

Shapiro, E. S. (1989). Academic skills problems: Direct assessment and intervention. New York, NY: The Guilford Press.

Shapiro, E. S., Hilt-Panahon, A., Gischlar, K. L., Semeniak, K., Leichman, E., & Bowles, S. (2012). An analysis of consistency between team decisions and reading assessment data

within an RTI model. *Remedial and Special Education, 33,* 335-347. Doi: 10.1177/0741932510397763

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized test: Readin and math performance in Pennsylvania. *Journal of Psychoeducational Assessment, 24,* 19-35. Doi10.1177/0734282905285237

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia.*Journal of Learning Disabilities, 25,* 618-629.doi:10.1016/j.cedpsych.2005.10.001

Silberglitt, B., Appleton, J. J., Burns, M., K., & Jimerson, S. R. (2006a). Examining the effects of grade retention on student reading performance: A longitudinal study. *Journal of School Psychology, 44,* 255-270.

Silberglitt, B. & Hintze, J. M. (2007). How much growth can we expect? A conditional analysis of R-CBM growth rates by level of performance. *Council for Exceptional Children, 74,* 71-84.

Silberglitt, B., Jimerson, S. R., Burns, M. K., & Appleton, J. J. (2006b). Does the timing of grade retention make a difference? Examining the effects of early versus later retention. *School Psychology Review, 35,* 134-141.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivation perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69,* 493-525. DOI: 10.1177/0013164408323233

Speece, D. L. (2002). Classification of learning disabilities: Convergence, expansion, and caution. In R. Bradley, L. Danielson, & D.P. Hallahan (Eds.), Identification of learning

disabilities: Research to practice (279-286). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Speece, D.L. & Ritchey, K. D (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, *38*, 387-399.

Stage, S. A., Sheppard, J., Davidson, M. M., & Browning, M. M. (2001).Prediction of first-graders' growth in oral reading fluency using kindergarten letter fluency. *Journal of School Psychology*, *39*, 225-237.

Stearns, E., Moller, S., Blau, J., & Potochnick, S. (2007). Staying back and dropping out: The relationship between grade retention and school dropout. *Sociology of Education, 80*, 210-240.DOI: 10.1177/003804070708000302

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. International Scholarly Research Notices, 2013. Retrieved from http://dx.doi.org/10.1155/2013/958530

Tindal, G., & Nese, J. F. T. (2013). *Oral reading fluency growth: a sample of methodology and findings*. Retrieved from the National Center on Assessment and Accountability for Special Education (NCAASE) website: http://ncaase.comn/publications/in-briefs

Tingle, L. R., Schoeneberger, J., & Algozzine, B. (2012). Does grade retention make a difference? *The Clearing House, 85*, 179-185.DOI: 10.1080/00098655.2012.679325

United States Department of Education. (2006). The early reading and mathematics achievement of children who repeated kindergarten or who began school a year late. *National Center for Education Statistics.*

United States Department of Education. (2011). Findings in brief: Reading and Mathematics 2011, National Assessment of Educational Progress at grades 4 and 8. *National Center for Education Statistics.*

United States Department of Education. (2013). Kindergarten entry status: On-time, delayed-entry, and repeating kindergartners. [Graphic illustration] *National Center for Education Statistics.* Retrieved from: http://nces.ed.gov/programs/coe/indicator_tea.asp

United States Department of Education. (2015a). Children and youth with disabilities. [Graphic illustration] *National Center for Education Statistics*. Retrieved from: http://nces.ed.gov/programs/coe/indicator_cgg.asp

United States Department of Education. (2015b). English language learners. [Graphic illustration] *National Center for Education Statistics*. Retrieved from: http://nces.ed.gov/programs/coe/indicator_cgf.asp

Urdan, T. S. (2010). *Statistics in plain English* (3rd ed.). New York: Routledge.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a Response to Interventio (RTI) model on identification of children for special education. *Journal of School Psychology, 45,* 225-256.

Vaughn, S., & Klinger, J. (2007). Overview of the three-tiered model of reading intervention. In D. Haager, J. Klinger, & S. Vaughn (Eds.), *Evidence-based reading practices for response to intervention* (3-10). Baltimore, Maryland: Paul H. Brookes Publishing Co.

Vaughn, S., Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (2007). Prevention and early identification of students with reading disabilities. In D. Haager, J. Klinger, & S. Vaughn (Eds.), *Evidence-based reading practices for response to intervention* (3-10). Baltimore, Maryland: Paul H. Brookes Publishing Co.

Vellutino, F. R., Sanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33,* 223-238.

Warren, J. R., & Saliba, J. (2012). First-through eighth-grade retention rates for all 50 states: A new method and initial results. *Educational Researcher, 41*, 320-329.DOI: 10.3102/0013189X12457813

Wu, W., West, S.G., & Hughes, J.N. (2008). Short-term effects of grade retention on the growth rate of Woodcock-Johnson III broad math and reading scores. *Journal of School Psychology,* 46. DOI:10.1016/j.jsp.2007.01.003

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumption of multiple regression: Correcting two misconceptions. *Practical Assessment, Research & Evaluation, 18.*

Zumeta, R. O., Compton, D. L., & Fuchs, L. S. (2012). Using word identification fluency to monitor first-grade reading development. *Exceptional Child, 78*, 201-220.